

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:09:35 ; Search time 22 Seconds
(without alignments)
5988.902 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagagag.....ttattataaaaaaaaaa 2239

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_090918_20329/app_query.fasta_1.2375
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_17_erunat_04042003_090918_20329 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMMap -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1156.5	28.1	515	2	US-09-073-362-3
2	1156.5	28.1	515	2	US-09-243-920-3
3	1131.5	27.5	511	2	US-09-073-362-1
4	1131.5	27.5	511	2	US-09-243-920-1
5	621	15.1	241	2	US-08-825-781-3
6	615	14.9	241	2	US-08-825-781-1
7	524.5	12.7	245	2	US-08-825-781-1
8	286	6.9	629	2	US-08-132-990A-8
9	286	6.9	629	5	PCT-US92-09382-8
10	282.5	6.9	521	4	US-09-134-001C-4290
11	279.5	6.8	622	2	US-08-132-990A-4
12	279.5	6.8	622	5	PCT-US92-09382-4

13	224.5	5.4	499	4	US-09-134-001C-5370	Sequence 5370, Ap
14	200.5	4.9	494	4	US-09-134-001C-4475	Sequence 4475, Ap
15	196	4.8	476	4	US-09-134-001C-3778	Sequence 3778, Ap
16	194	4.7	465	4	US-09-134-001C-4428	Sequence 4428, Ap
17	192.5	4.7	367	2	US-08-132-990A-2	Sequence 2, Appli
18	192.5	4.7	367	5	PCT-US92-09382-2	Sequence 2, Appli
19	190.5	4.6	456	4	US-09-134-001C-3771	Sequence 3771, Ap
20	178	4.3	344	4	US-09-134-001C-4175	Sequence 4175, Ap
21	169.5	4.1	503	4	US-09-068-195-24	Sequence 24, Appl
22	166	4.0	453	2	US-08-132-990A-6	Sequence 6, Appli
23	166	4.0	453	5	PCT-US92-09382-6	Sequence 6, Appli
24	163.5	4.0	482	4	US-09-134-001C-4309	Sequence 4309, Ap
25	158.5	3.8	414	4	US-09-134-001C-3357	Sequence 3357, Ap
26	158	3.8	496	4	US-09-134-001C-3001	Sequence 3001, Ap
27	141	3.4	1212	4	US-09-268-866-2	Sequence 2, Appli
28	130.5	3.2	165	4	US-08-858-207A-462	Sequence 462, App
29	129	3.1	249	4	US-09-134-001C-5030	Sequence 5030, Ap
30	129	3.1	430	4	US-09-134-001C-2981	Sequence 2981, Ap
31	126	3.1	454	4	US-09-134-001C-3501	Sequence 3501, Ap
32	123.5	3.0	432	4	US-09-370-253-12	Sequence 12, Appl
33	123.5	3.0	527	4	US-09-134-001C-3358	Sequence 3358, Ap
34	119.5	2.9	3177	2	US-08-477-451-4	Sequence 4, Appli
35	117	2.8	349	4	US-09-134-001C-4519	Sequence 4519, Ap
36	115.5	2.8	554	4	US-09-134-001C-5109	Sequence 5109, Ap
37	112	2.7	473	4	US-09-637-118B-2	Sequence 2, Appli
38	111.5	2.7	413	4	US-09-134-001C-3702	Sequence 3702, Ap
39	109	2.6	525	4	US-09-351-224E-8	Sequence 8, Appli
40	108	2.6	599	1	US-07-879-617A-10	Sequence 10, Appl
41	108	2.6	599	1	US-08-295-814E-11	Sequence 11, Appl
42	108	2.6	599	1	US-08-240-783B-4	Sequence 4, Appli
43	108	2.6	599	1	US-08-753-985-10	Sequence 10, Appl
44	108	2.6	599	3	US-09-084-813-4	Sequence 4, Appli
45	108	2.6	599	4	US-09-343-361-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-09-073-362-3
; Sequence 3, Application US/09073362
; Patent No. 5942399
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166

US-00-667-170A-440 (1-2239) X US-09-073-362-3 (1-515)

```

QY      1114 GAGCTGCTGCTTTCAAATGCAGTGGCAGCTACTTTTCTGAGCGCTACTGGGAATTTC 1173
       ::::||||| |||:::||||| ||||| ||||| ::::: ||| |||
Db      297 AspValLeuSerSerAspAlaValAlaValThrPheAlaAspGlnThrPheGlyMetPhe 316

QY      1174 TCATTAGCAGTCCGATCTTTGTGCCCTCTCTGCTTTGGCTCCATGAACGGTGGTG 1233
       ||| :::||||| ||||| ||||| ||||| ::||| :::
Db      317 SerTrpThrIleProIleAlaValAlaLeuSerCysPheGlyGlyLeuAsnAlaSerIle 336

QY      1234 TTTGCTGTCTCCAGETATTCTATGTTGCGTCTCGAGGGTCACCTTCCAGAAATCCTC 1293
       ||||| ||||| ||||| ::||| ||||| ||||| ||||| :::::
Db      337 PheAlaSerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProAspLeuLeu 356

QY      1294 TCCATGATTTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTTGACCCCTTTG 1353
       ||||| ||||| ::| ||||| ::||| ||||| :::::
Db      357 SerMetIleHisIleGluArgPheThrProIleProAlaLeuLeuPheAsnCysThrMet 376

QY      1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTTTTGAATTTCCCTAGTTTGCC 1413
       ::||| ||| ||||| ||||| ||||| ||||| ||||| :::::
Db      377 AlaIleIleTyrLeuIleValGluAspValPheGlnLeuIleAsnTyrPheSerPheSer 396

QY      1414 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473
       ||| |||:::|||||::||| ||| :::|||||:::||| |||
Db      397 TyrTrpPhePheValGlyLeuSerValValGlyGlnLeuTyrLeuArgTrpLysGluPro 416

QY      1474 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCCAGCTTTGTTTTCTTCACATGC 1533
       ||||| ||||| ||||| ::||| ||| ::|||
Db      417 LysArgProArgProLeuLysLeuSerValPhePheProIleValPheCysIleCysSer 436

QY      1534 CTCCTCATGGTGGCCCTTTCCCTCTATTCCGGACCCCATTTAGTACAGGGATTGGCTTCGTC 1593
       ::|||::||| ::| |||:::|||| |::::| |||||
Db      437 ValPheLeuValIleValProLeuPheThrAspThrIleAsnSerLeuIleGlyIleGly 456

QY      1594 ATCACTCTGACTGGAGTCCCTCGGTATATCTC-----TTTATTATATGGACAAGAAA 1647
       ||| |||:::||||| |||::: |||::: |||:::
Db      457 IleAlaLeuSerGlyValProPheTyrPheMetGlyValTyrLeuProGluSerArgArg 476

QY      1648 CCCAGGTGGTTAGAATAAATGTCCAGAGAAATAAACAGA---ACATTACAAATAATACTG 1704
       ||| ::| ||| ::| ||||| ||| |||||
Db      477 ProLeuPheIleArgAsnValLeuAlaAlaIleThrArgGlyThrGlnGlnLeuCysPhe 496

QY      1705 GAAGTTGTACCAGAGAAGAT 1725
       |||::| ||| |||
Db      497 CysValLeuThrGluLeuAsp 503

RESULT 2
US-09-243-920-3
; Sequence 3, Application US/09243920
; Patent No. 5981242
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,920
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/073,362

```

ИТС-09-243-920-3

US-09-243-920-3
; Sequence 3, Application US/09243920
: Patent No. 5981242

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Yue, Henry

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOGY

NUMBER OF SEQUENCES:

CORRESPONDENCE ADDRESS:

ADDRESS: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY. Palo Alto

CITY: LOS ANGELES
STATE: CA

COUNTRY. U

COUNTRY: GERM
7TP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

COMPUTER: IBM COMP
OPERATING SYSTEM: DOS

OPERATING SYSTEM: FastSEO for
SOFTWARE: FastSEO for

CURRENT APPLICATION DATA:

ADDITIONAL INFORMATION

APPLICATION NUMBER:
FILING DATE:

FILED DATE:
CLASSIFICATION:

CLASSIFICATION:
PRIOR APPLICATION DATA

APPLICANT NUMBER: 09/073,362


```

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0514 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 511 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ESOGTUT02
CLONE: 2667831

```

Alignment Scores:		
pred. No.:	2.85e-110	511
Score:	1131.50	214
Percent Similarity:	68.00%	109
Best Local Similarity:	45.05%	149
Query Match:	27.46%	3
DB:	2	2
Length:		
Matches:		
Conservative:		
Mismatches:		
Indels:		
Gaps:		

1708-440 (1-2239) x US-09-073-362-1 (1-511)

	US-09-687-170R	440	(220)	
QY	310	GGAAGACGCGCTTTTCAGGAAGAAGTGCAGCTGAAAGCAATCCTACTTG	369		
Ddb	21	GlyAspGlyAlaSerProGluValLysLeuIlePheValSerLeuLeu	40		
QY	370	AGGGAGGTCTCCATTATCATNTGGACCATTGGAGCACGAATACTTCATCTCTCCTTAAG	429		
Ddb	41	AsnGlyValCysLeuIleValGlyAsnMetileGlySerGlyIlePheValSerProLys	60		
QY	430	GGCGTGCTCCAGAACACGGGCAGCGTGGGCATGTCTCTGCCAATCTGGACGGTGTGTGGG	489		
Ddb	61	GlyValLeuIleTyrrSerAlaSerPheGlyLeuSerLeuValIleTrpAlaValGlyGly	80		
QY	490	GTCCTGTGCATPATTTGGAGCTTTGTCTTATGCTGAATTTGGACAACACTATAAAGAAATCT	549		
Ddb	81	LeuPheSerValPheGlyAlaLeuCystyrAlaGluleuGlyThrthrIleLysLysSer	100		
QY	550	GGAGGTCAATTACACATATATTTTTGGAAAGTCTTTTGGTCATTACCCAGCTTTTGTACGAGTC	609		
Ddb	101	GlyAlaSerTyrrAlaTyrrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu	120		
QY	610	TGGGTGGAACCTCCATCATAATACGCCCTGCAGCTACTGCTGTGATATCCCCTGGCATTTTGA	669		
Ddb	121	TriprThrSerLeuLeuIlelleGluProThrSerGlnAlaIlelleAlaIleThrpPheAla	140		
QY	670	CGCTACATTTCTGGAACCATTTTTTTATTCATGTGAAATCCCTGAACTTGGATCAAGCTC	729		
Ddb	141	AsnTyrrMetValGlnProLeuPheProSerCysPheAlaProTyrrAlaAlaSerArgLeu	160		
QY	730	ATTACAGCTGTGGGCATAACTGTAGTGTGCTTAATAGCATGAGTGTACGCTGGAGC	789		
Ddb	161	LeuAlaAlaAlaCysilleCysLeuLeuThrPheilleAsnCysAlatyrValltyrTpGly	180		
QY	790	GCCCCGATCCAGATTTTCTTTAACCTTTTGCRAAGCTCACAGCAAFTTCTGATAATATAGTC	849		
Ddb	181	ThrLeuValGlnAspilePhethrTyrrAlaLysValLeuAlaLeuilleAlavalleval	200		

QY	850	CCTGGAGTTATGCAGCTAATTAAAGGTCAAACCGAGAACA	TTTTAAAGACGCCCTTTTCAGGA	909
Dd	201	AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly	220	
QY	910	AGAGATTCAAGTATTACGCGGTGCCACTGGCTTTTATTATGGAATGTATGCATATGCT	969	
Dd	221	SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer	240	
QY	970	GGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCCTGAAAAACCATTCCTCC	1029	
Dd	241	GlyTrpAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro	260	
QY	1030	CTTGCAATATGTAATCCATGGCCATGTGCACCATTGGCTATGTGCTGACAAATGTGGCC	1089	
Dd	261	LeuSerIleGlyIleSerMetProIleValThrIleIleTyrIleLeuThrAsnValAla	280	
QY	1090	TACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTTCAAAATGCAGTGGCAGTCACTTT	1149	
Dd	281	TyrTyrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe	300	
QY	1150	TCTGACGGCTACTGGGAAATTTCTCATTAAGCAGTTCCGATCTTGTGTCCTCTCCTGC	1209	
Dd	301	AlaAspGlnIlePheGlyIlePheAsnTrpIleProLeuSerValAlaLeuSerCys	320	
QY	1210	TTTGGCTCCATGAACGGTGGTGTGTTTGTGTCCTCCAGGTTATTCTATGTTGCGTCTCGA	1269	
Dd	321	PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeupheValGlySerArg	340	
QY	1270	GAGGTCACCTTCCAGAAATCTCTCCATGATTCGCGAACGACACACTCCTCTACCA	1329	
Dd	341	GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro	360	
QY	1330	GCTGTATTGTTTGCACCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTCGACAGT	1389	
Dd	361	SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln	380	
QY	1390	CTTTTGAATTCCTCAGTTTTTGGCAGGTGGCTTTTATGGGCTGGCAGTGTCTGGGCTG	1449	
Dd	381	LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln	400	
QY	1450	ATTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCACTGTTCATC	1509	
Dd	401	LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe	420	
QY	1510	CCAGCTTGTGTTTCCCTTCACATGCCCTCTTCATGGTTGCCCTTCCCTCTATTCGACCCA	1569	
Dd	421	ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr	440	
QY	1570	TTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTCGCTATTATCTCTTT	1629	
Dd	441	IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle	460	
QY	1630	ATT-----ATATGGGACAAGAAACCCAGGTGGTTTAGAATAATGTCAGAGAAAATAACC	1683	
Dd	461	IleArgValProGluHisLysArgProLeuTyrLeuArgArgIleValGlySerAlaThr	480	
QY	1684	AGAACATTACAAATAATA--CTGGAAGTTGTACCCAGAAGAAGAT	1725	
QY				

RESULT 4

US-09-243-920-1

US-09-243-920-1 Application US/09243920

Sequence I, Application No. 5981242

Patent No. 3961242
GENERAL INFORMATION:

GENERAL INFORMATION: Hillma

APPLICANT: Yue, Henry

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: AMINO

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc

Thu Apr 17 09:48:24 2003

GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN TUMOR-ASSOCIATED MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/825,781
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0262 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 392990
US-08-825-781-3

Alignment Scores:
Pred. No.: 1e-56
Score: 621.00
Percent Similarity: 68.49%
Best Local Similarity: 47.90%
Query Match: 15.07%
DB: 2

US-09-667-170A-440 (1-2239) x US-08-825-781-3 (1-241)

QY 1009 AACCCCTGAAAAAACCATTCCTCCCTTGCAATATGTATATCCATGGCCATTGTCACCATTTGGC 1068
Db 3 AsnProTyrArgAsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuVal 22
QY 1069 TATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCA 1128
Db 23 TyrValLeuThrAsnLeuAlaTyrPheThrThrLeuSerThrAsnGlnMetLeuThrSer 42
QY 1129 AATGAGTGGCAGTACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCG 1188
Db 43 GluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIlePro 62
QY 1189 ATCTTTGTTGCCCTCTCCTGCTCCATGAACGGTGGTGTGTGTGCTGTCTCCAGG 1248
Db 63 ValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThrSerArg 82
QY 1249 TTATTCATGTTGGTCTCGAGAGGGTCACCTTCAGAAATCCTCTCCATGATTATGTC 1308
Db 83 LeuPhePheValGlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisPro 102

QY 1309 CGAAGCACACTCCTCTACAGCTGTTATTGTTTGGACCCCTTTGACAATGATGATGCTC 1368
Db 103 GlnLeuLeuThrProValProSerLeuValPheThrCysValMetThrLeuMetTyrAla 122
QY 1369 TTCTCTGGAGACCTCGACAGTCTTTGAATTTCCACAGTTTGGCCAGGTGGCTTTTATT 1428
Db 123 PheSerArgAspIlePheSerIleIleAsnPhePheSerPheAsnTrpLeuCysVal 142
QY 1429 GGGCTGGCAGTGTGGCTGATTATCTTCGATACAAATGCCAGATATGATGCTGCTCCT 1488
Db 143 AlaLeuAlaIleIleGlyMetMetTrpLeuArgPheLysLysProGluLeuGluArgPro 162
QY 1489 TTCAAGGTGCCACTGTTTCATCCAGCTTGTGTTTTCCTTCACATGCCCTCTTCATGTTGCC 1548
Db 163 IleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAla 182
QY 1549 CTTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCATCTGACTGGA 1608
Db 183 ValSerPheTrpLysThrProLeuGluCysGlyIleGlyPheAlaIleIleLeuSerGly 202
QY 1609 GTCCCTCGTATTATCTCTTTTATTATATGGACAAGAAACCCAGTGGTGTAGATAATG 1668
Db 203 LeuProValTyrPhePheGlyValTrpTrpLysAsnLysProLysTrpIleLeuGlnVal 222
QY 1669 TCAGAGAAAATAACGAGAACATTACAAATAATACCTGGAAGTTGTACAGAGAA 1722
Db 223 IlePheSerValThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 240

RESULT 6

US-08-825-781-4
Sequence 4, Application US/08825781
Patent No. 5843727

GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN TUMOR-ASSOCIATED MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/825,781
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0262 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 181908

US-08-825-781-4

Alignment Scores:

Pred. No.: 4.29e-56 Length: 241
Score: 615.00 Matches: 115
Percent Similarity: 68.07% Conservative: 47
Best Local Similarity: 48.32% Mismatches: 76
Query Match: 14.93% Indels: 0
DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x US-08-825-781-4 (1-241)

QY 1009 AACCTGAAACCAATCCCTTGCAATATGATATATCCATGGCCATTGTCACCATGGC 1068
Db 3 AsnProTyrArgAsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuVal 22
QY 1069 TATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCA 1128
Db 23 TyrValLeuThrAsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSer 42
QY 1129 AATGAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCG 1188
Db 43 GluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIlePro 62
QY 1189 ATCTTTGTTGCCCTCTCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGCTCCAGG 1248
Db 63 ValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArg 82
QY 1249 TTATTTCTATGTTGGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTC 1308
Db 83 LeuPhePheValGlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisPro 102
QY 1309 CGCAAGCACACTCTCTACCACTGTTATTTGTTGACCCCTTTTGACATGATAATGCTC 1368
Db 103 GlnLeuLeuThrProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAla 122
QY 1369 TTCTCTGGAGACCTCGACAGCTTTTGAATTTCTCAGTTTGGCCAGTGGCTTTTATT 1428
Db 123 PheSerLysAspIlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysVal 142
QY 1429 GGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCT 1488
Db 143 AlaLeuAlaIleIleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgPro 162
QY 1489 TTCAAGTGGCCACTGTTTCATCCAGCTTTGTTTCTTCCATGCTCTTTCATGGTGGC 1548
Db 163 IleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAla 182
QY 1549 CTTTCCCTCTATTCCGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGA 1608
Db 183 ValSerPheTrpLysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGly 202
QY 1609 GTCCCTGCGTATTATCTCTTATTATATGGGACAAAGAACCCAGGTGGTTTGAATAATG 1668
Db 203 LeuProValTyrPhePheGlyValTrpTrpLysAsnLysProLysTrpLeuGlnGly 222
QY 1669 TCAGAGAAATACCAAGAACATTACAAATAATACTGGAAGTTGTACCAGAGAA 1722
Db 223 IlePheSerThrThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 240

RESULT 7

US-08-825-781-1

; Sequence 1, Application US/08825781
; Patent No. 5843727

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Shah, Purvi

; TITLE OF INVENTION: HUMAN TUMOR-ASSOCIATED MEMBRANE PROTEIN

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/825,781
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0262 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PANTUT02
; CLONE: 2236771
; US-08-825-781-1

Alignment Scores:
Pred. No.: 1.44e-46 Length: 245
Score: 524.50 Matches: 102
Percent Similarity: 67.25% Conservative: 52
Best Local Similarity: 44.54% Mismatches: 72
Query Match: 12.73% Indels: 3
DB: 2 Gaps: 2

US-09-667-170A-440 (1-2239) x US-08-825-781-1 (1-245)

QY 1048 ATGGCCATTGTCACCATTTGGCTATGTGCTGACAAATGCGCTACTTTACGACCATTAAT 1107
Db 1 MetProIleValThrIleIleTyrIleLeuThrAsnValAlaTyrTyrThrValLeuAsp 20
QY 1108 GCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGA 1167
Db 21 MetArgAspIleLeuAlaSerAspAlaValAlaValThrPheAlaAspGlnIlePheGly 40
QY 1168 AATTTCTCATAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGT 1227
Db 41 IlePheAsnTrpIleIleProLeuSerValAlaLeuSerCysPheGlyGlyLeuAsnAla 60
QY 1228 GGTGTGTTGCTGCTCCAGGTTATTTCTATGTTGCGTCTCGAGAGGGTCCACCTTCCAGAA 1287
Db 61 SerIleValAlaAlaSerArgLeuPhePheValGlySerArgGluGlyHisLeuProAsp 80
QY 1288 ATCCTCTCCATGATTCATGTCGCAAGCACACTCTCTACCACTGCTTATTGTTTGCAC 1347
Db 81 AlaIleCysMetIleHisValGluArgPheThrProValProSerLeuLeuPheAsnGly 100
QY 1348 CCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTGAATTTCCTCAGT 1407
Db 101 IleMetAlaLeuIleTyrLeuCysValGluAspIlePheGlnLeuIleAsnTyrTyrSer 120
QY 1408 TTGCCCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAA 1467
Db 121 PheSerTyrTrpPhePheValGlyLeuSerIleValGlyGlnLeuTyrLeuArgTrpLys 140
QY 1468 TGCCCAGATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCTC 1527

us-09-667-170a-440.n2p.ra1

Thu Apr 17 09:48:24 2003

Db 141 GluProAspArgProArgProLeuLysLeuSerValPhePheProIleValPheCysLeu 160
QY 1528 ACATGCTCTTTCATGGTGGCTTCCCTCTATTCGGACCCCATTTTAGTACAGGGATTGGC 1587
Db 161 CysThrIlePheLeuValAlaValProLeuTyrSerAspThrIleAsnSerLeuIleGly 180
QY 1588 TTCGTCACTCACTGACTGGAGTCCCTGCGTATTATCTCTTTATT-----ATATGGGAC 1641
Db 181 IleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIleIleArgValProGluHis 200
QY 1642 AAGAAACCCAGGTGGTTTAGAATAATGTCAGAGAAATAACCAAGACATTACAAATAATA 1701
Db 201 LysArgProLeuTyrLeuArgArgIleValGlySerAlaThrArgTyrLeuGlnValLeu 220
QY 1702 ---CTGGAAGTGTACCAGAGAGAT 1725
Db 221 CysMetSerValAlaAlaGluMetAsp 229

RESULT 8
US-08-132-990A-8
; Sequence 8, Application US/08132990A
; Patent No. 5834589
; GENERAL INFORMATION:
; APPLICANT: MERUELO, DANIEL
; APPLICANT: YOSHIMOTO, TAKAYUKI
; TITLE OF INVENTION: Human Retrovirus Receptor and DNA Coding Therefor
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/132,990A
; FILING DATE: 07-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/084,729
; FILING DATE: 29-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05569
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/899,075
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/806,178
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/627,950
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8105-004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 629 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

Alignment Scores: 2.93e-21 Length: 629
Pred. No.: 286.00 Matches: 103
Score: 43.16% Conservative: 80
Percent Similarity: 24.29% Mismatches: 183
Best Local Similarity: 6.94% Indels: 58
Query Match: 2 Gaps: 12
Db: 2
US-09-667-170A-440 (1-2239) x US-08-132-990A-8 (1-629)
QY 331 AGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTTACTGAGGGGAGTCTCCATTATCAT 390
Db 23 ArgGluGluThrArgLeuSerArgCysLeuAsnThrPheAspLeuValAlaLeuGlyVal 42
QY 391 GGCACCATCATTTGGAGCAGGAATCTTCACTCTCTCCTAAGGGCGTCTCCAGAACACGGGC 450
Db 43 GlySerThrLeuGlyAlaGlyValTyrValLeuAlaGlyAlaValAlaArgGlu----- 60
QY 451 AGCGTGGGCATGTCTCTGACCATC---TGGACGGTGTGTGGGGTCTGTCTACTATTTGGA 507
Db 61 AsnAlaGlyProAlaIleValIleSerPheLeuIleAlaAlaLeuAlaSerValLeuAla 80
QY 508 GCTTTGTCTTATGCTGAATTTGGGAACAACCTATAAAGAAATCTGGAGGTCATTACACATAT 567
Db 81 GlyLeuCysTyrGlyGluPheGlyAlaArgValProLysThrGlySerAlaTyrLeuTyr 100
QY 568 ATTTTGGAGTCTTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGGACTCTCTC--- 624
Db 101 SerTyrValThrValGlyGluLeuTrpAlaPheIleThrGlyTrpAsnLeuIleLeuSer 120
QY 625 ---ATAATACGCCCTGCAGCTACTGCT---GTGATATCCCTGGCATTTTGGACGCTACATT 678
Db 121 TyrIleIleGlyThrSerSerValAlaAlaArgAlaTrpSerAlaThrPheAspGluLeuIle 140
QY 679 CTGGAACCATTT-----TTTATTCAA 699
Db 141 GlyArgProIleGlyGluPheSerArgThrHisMetThrLeuAsnAlaProGlyValLeu 160
QY 700 TGTGAAATCCCTGAACTTGGCATCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGATG 759
Db 161 AlaGluAsnProAspIlePheAlaValIleIle-----IleLeuIleLeuThr 176
QY 760 GTCCTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTCTTAACTTTTGC 819
Db 177 GlyLeuLeuThrLeuGlyValLysGluSerAlaMetValAsnLysIlePheThrCysIle 196
QY 820 AAGCTCACAGCAATTTCTATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAAGGTCAA 879
Db 197 AsnValLeuValLeuGlyPheIleMetValSerGly-----PheValLysGlySer 213
QY 880 ACGCAGAACTTTAAA----- 894
Db 214 ValLysAsnTrpGlnLeuThrGluGluAspPheGlyAsnThrSerGlyArgLeuCysLeu 233
QY 895 -----GACGCCCTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTAT 948
Db 234 AsnAsnAspThrLysGluGlyLysProGlyValGlyPheMetPropheGlyPheSer 253
QY 949 TATGGAATG-----TATGCATATGCTGGTGGTTTACCTCAAC 987
Db 254 GlyValLeuSerGlyAlaAlaThrCysPheTyrAlaPheValGlyPheAspCysIleAla 273
QY 988 TTTGTTACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTTCCCTTGCATATATGATATCC 1047
Db 274 ThrThrGlyGluGluValLysAsnProGlnLysAlaIleProValGlyIleValAlaSer 293
QY 1048 ATGGCCATTGTACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCAATAAT 1107
Db 294 LeuLeuIleCysPheIleAlaTyrPhe-----GlyValSerAlaAlaLeuThrLeuMet 311
QY 1108 GCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGA 1167

Db 312 MetProTyrPheCysLeuAspAsnSerProLeuProAspAlaPheLysHisValGly 331
 QY 1168 -----AATTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGC 1215
 Db 332 TrpGluGlyAlaLysTyrAlaValAlaValGlySerLeuCysAlaLeuSer-----Ala 349
 QY 1216 TCCATGAACGGTGGTGTGTTGCTGCTCTCCAGGTTATCTATGTTGGCTCGAGAGGGT 1275
 Db 350 SerLeuLeuGlySerMetPheProMetProArgValIleTyrAlaMetAlaGluAspGly 369
 QY 1276 CACCTTCCAGAAATCTCTCCATGATTCATGTCCGCAAGCACACTCCTCTACAGCTGTT 1335
 Db 370 LeuLeuPheLysPheLeuAlaAsnValAsnAspArgThrLysThrProIleIleAlaThr 389
 QY 1336 ATTGTTTTCACCCCTTTTGACAAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTG 1395
 Db 390 LeuAlaSerGlyAlaValAlaAlaValMetAlaPheLeuPheAspLeuLysAspLeuVal 409
 QY 1396 AATTCTCCTCAGTTTGGCCAGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTAT 1455
 Db 410 AspLeuMetSerIleGlyThrLeuLeuAlaTyrSerLeuValAlaAlaCysValLeuVal 429
 QY 1456 CTTGCATACAAA 1467
 Db 430 LeuArgTyrGln 433

RESULT 9

PCT-US92-09382-8
 ; Sequence 8, Application PC/TUS9209382
 ; GENERAL INFORMATION:
 ; APPLICANT: MERUELO, DANIEL
 ; TITLE OF INVENTION: YOSHIMOTO, TAKAYUKI
 ; TITLE OF INVENTION: Human Retrovirus Receptor and DNA Coding
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street, N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/09382
 ; FILING DATE: 19921213
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Livnat, Shmuel
 ; REGISTRATION NUMBER: 33,949
 ; REFERENCE/DOCKET NUMBER: MERUELO=1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202 628-5197
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 629 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US92-09382-8

Alignment Scores:
 Pred. No.: 2.93e-21
 Score: 286.00
 Percent Similarity: 43.16%
 Best Local Similarity: 24.29%
 Query Match: 6.94%
 DB: 5
 Length: 629
 Matches: 103
 Conservative: 80
 Mismatches: 183
 Indels: 58
 Gaps: 12

US-09-667-170A-440 (1-2239) x PCT-US92-09382-8 (1-629)
 QY 331 AGAGAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCAT 390
 Db 23 ArgGluGluThrArgLeuSerArgCysLeuAsnThrPheAspLeuValAlaLeuGlyVal 42
 QY 391 GGCACCATCATTTGGAGCAGGAATCTTCACTCTCTCTAAGGGCGTCTCCAGAACACGGGC 450
 Db 43 GlySerThrLeuGlyAlaGlyValTyrValLeuAlaGlyAlaValAlaArgGlu----- 60
 QY 451 AGCGTGGGCATGCTCTGACCATC---TGGACGGTGTGTGGGCTGTGTCACTATTGGA 507
 Db 61 AsnAlaGlyProAlaIleValIleSerPheLeuIleAlaAlaLeuAlaSerValLeuAla 80
 QY 508 GCTTGTCTTATGCTGAATTTGGAAACAATATAAGAAATCTGGAGGTCACTACACATAT 567
 Db 81 GlyLeuCysTyrGlyGluPheGlyAlaArgValProLysThrGlySerAlaTyrLeuTyr 100
 QY 568 ATTTTGAAGTCTTTGGTCCATTACCAAGTCTTGTACGAGTCTGGGTGGAACCTCCTC--- 624
 Db 101 SerTyrValThrValGlyGluLeuTrpAlaPheIleThrGlyTrpAsnLeuIleLeuSer 120
 QY 625 ---ATAATACGCCCTGCAGCTACTGCT---GTGATATCCTCGGATTTGGACGTACATT 678
 Db 121 TyrIleIleGlyThrSerSerValAlaArgAlaTrpSerAlaThrPheAspGluLeuIle 140
 QY 679 CTGGAACCATTT-----TTTATTCAA 699
 Db 141 GlyArgProIleGlyGluPheSerArgThrHisMetThrLeuAsnAlaProGlyValLeu 160
 QY 700 TGTGAATCCCTGAACCTTGCAGTCAAGCTCACTACAGCTGTGGGCATACACTGTAGTGATG 759
 Db 161 AlaGluAsnProAspIlePheAlaValIleIle-----IleLeuIleLeuThr 176
 QY 760 GTCCTAAATAGCATGAGTGTACAGCTGGAGCGCCGCGATCCAGATTTTCTTAACCTTTTGC 819
 Db 177 GlyLeuLeuThrLeuGlyValLysGluSerAlaMetValAsnLysIlePheThrCysIle 196
 QY 820 AAGCTCACAGCAATCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAA 879
 Db 197 AsnValLeuValLeuGlyPheIleMetValSerGly-----PheValLysGlySer 213
 QY 880 ACGCAGAACTTTAA-----894
 Db 214 ValLysAsnTrpGlnLeuThrGluGluAspPheGlyAsnThrSerGlyArgLeuCysLeu 233
 QY 895 -----GACGCCCTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTAT 948
 Db 234 AsnAsnAspThrLysGluGlyLysProGlyValGlyPheMetPropheGlyPheSer 253
 QY 949 TATGGAATG-----TATGCATATGCTGGTGGTGGTTTACCTCAAC 987
 Db 254 GlyValLeuSerGlyAlaAlaThrCysPheTyrAlaPheValGlyPheAspCysIleAla 273
 QY 988 TTTGTTACTGAAGAAGTAGAAACCCCTGAAAAAACCACTCCCTTGCATATATGATATCC 1047
 Db 274 ThrThrGlyGluGluValLysAsnProGlnLysAlaIleProValGlyIleValAlaSer 293
 QY 1048 ATGGCCATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTAGCACCAATAAT 1107
 Db 294 LeuLeuIleCysPheIleAlaTyrPhe-----GlyValSerAlaAlaLeuThrLeuMet 311
 QY 1108 GCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGTGCACCTTTTCTGAGCGGCTACTGGA 1167
 Db 312 MetProTyrPheCysLeuAspAsnSerProLeuProAspAlaPheLysHisValGly 331
 QY 1168 -----AATTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGC 1215
 Db 332 TrpGluGlyAlaLysTyrAlaValAlaValGlySerLeuCysAlaLeuSer-----Ala 349
 QY 1216 TCCATGAACGGTGGTGTGTTGCTGCTCTCCAGGTTATCTATGTTGCTGCGTCTCGAGAGGGT 1275
 Db 350 SerLeuLeuGlySerMetPheProMetProArgValIleTyrAlaMetAlaGluAspGly 369

QY 1276 CACCTTCCAGAAATCCTCTCCATGATTTCATGTCGCGAAGCACACTCCTCTACCAGCTGTT 1335
Db 370 LeuLeuPheLysPheLeuAlaAsnValAsnAspArgThrLysThrProIleAlaThr 389
QY 1336 ATTGTTTTCACCCCTTTTGACAAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTT 1395
Db 390 LeuAlaSerGlyAlaValAlaValMetAlaPheLeuPheAspLeuLysAspLeuVal 409
QY 1396 AATTTCCTCAGTTTTCAGAGGTGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTAT 1455
Db 410 AspLeuMetSerIleGlyThrLeuLeuAlaTyrSerLeuValAlaAlaCysValLeuVal 429
QY 1456 CTTCGATACAAA 1467
Db 430 LeuArgTyrGln 433

RESULT 10

US-09-134-001C-4290
; Sequence 4290, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4290
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4290

Alignment Scores:
Pred. No.: 6.2e-21 Length: 521
Score: 282.50 Matches: 135
Percent Similarity: 43.39% Conservative: 108
Best Local Similarity: 24.11% Mismatches: 206
Query Match: 6.86% Indels: 111
DB: 4 Gaps: 26

US-09-667-170A-440 (1-2239) x US-09-134-001C-4290 (1-521)

QY 127 TTTTAT---CATATCTGGATTTTGGATTTCTTTGTTTCTCATCACTGGATTCAGGA 183
Db 8 PheTyrLeuHisValAspPheLeuIleArgIleValLysGluCysLeuLeuThr--- 26
QY 184 AAGCCTGTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGGAATGTTAACGGGAGG 243
Db 27 ---HisValAsnSerLysLeuMetLeuTrpArgPheValMetGlySerPhePheAsnArg 45
QY 244 CTGCCTTCCCTGGGCAACAAGGAGCCACCTGGCAGGACGCTTTTCAGGAAGAGACGCC 303
Db 46 MetThr-----ArgLysGluAsnPro-----ThrIle 54
QY 304 TTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAAAGTGCAGCTGAGAGGAAAGTCACT 363
Db 55 TyrGlnSerLysAspGly-----HisLeuLysArgThrLeuArg 67
QY 364 TTTACTAGGGGAGTCTCCATTATCATTTGGCACCATTATGGCAGGAGATCTTCTCTCT 423
Db 68 ValArgAspPheLeuAlaLeuGlyValGlyThrIleValSerThrSerIlePheThrLeu 87
QY 424 CCTAAGGGCGTGTCTCCAGAACACGGGACGCTGGGCGATCTCTGACCATCTGGACGGTG 483
Db 88 ProGlyValValAlaAlaGluHisAlaGlyProAlaValSerLeuSer---PheLeuLeu 106

QY 484 TGTGGGGTCTGTCTACTATTGGAGCTTTGTCTTATGCTGAATTGGGAACAACATAAAG 543
Db 107 AlaAlaIleValAlaGlyLeuValAlaPheThrTyrAlaGluMetAlaSerThrMetPro 126
QY 544 AATCTGGAGTCACTACACATATATTGGAAGCTTTGGTCCATTACCAGCTTTTGTGA 603
Db 127 PheAlaGlySerAlaTyrSerTrpIleAsnValLeuPheGlyGluLeuPheGlyTrpVal 146
QY 604 CGAGTCTGGTGGAACTCCTCATATAATACGCCCTGCAGCTACTGTGTATATCCTGGCA 663
Db 147 AlaGlyTrpAla---LeuLeuAlaGluTyrPheIleAlaValAlaPheValAlaSerGly 165
QY 664 TTTTGA-----CGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAA 714
Db 166 PheSerAlaAsnLeuArgGlyLeuIleAlaProLeuGlyIle-----SerLeuProLys 183
QY 715 CTT-----GCGATCAAGCTCATTACAGCTGTGGGC 744
Db 184 SerLeuSerAsnProPheGlySerAsnGlyGlyValIleAspIleIleAlaValVal 203
QY 745 ATAAGTGTAGTGTGCTCTAAATAGCATGAGTGTGAGCTGGAGCGCGGATCCAGATT 804
Db 204 IleIleLeuThrAlaLeuLeuSerArgGlyMetAsnGluAlaAlaArgMetGluAsn 223
QY 805 TTCTTAACCTTTTGCAGCTCAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAG 864
Db 224 ValLeuValIleLeuLysValLeuAlaIleIleLeuPheValIleValGlyLeuThrAla 243
QY 865 CTAATTTAAAGTCAAACGCGAGAACTTTAAAGAC----- 897
Db 244 Ile-----AsnPheSerAsnTyrIleProPheIleProGluHisLys 257
QY 898 -----GCCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCT 942
Db 258 ValThrGluThrGlyAspPheGlyGly----- 266
QY 943 TTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT----- 993
Db 267 ---TrpGlnGlyIleTyrAlaGlyValSerMetIlePheLeuAlaTyrIleGlyPheAsp 285
QY 994 -----ACTGAAGAAGTAGAAACCCCTGAAACACCATTTCCCTTGCATAA 1038
Db 286 SerIleAlaAlaAsnSerAlaGluAlaIleAsnProGlnLysThrMetProArgGlyIle 305
QY 1039 TGTATATCCATGGCCATT-----GTCACCATTTGGCTATGTGTGACAAAT 1083
Db 306 LeuGlySerLeuIleValAlaIleValLeuPheValAlaValAlaLeuValGly 325
QY 1084 GTGGCTTACTTTACGACCATTAATGCTGTGAGGAGCTGTGCTTTCAATGCAGTGGCAGTG 1143
Db 326 MetPheHisTyrSerGln-----TyrAlaAspAsnAlaGluProVal 339
QY 1144 ACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCA-----GTTCCGATCTTTGTT 1197
Db 340 GlyTrpAlaLeuArgGluSerGlyHisGlyIleIleAlaAlaIleValGlnAlaIleSer 359
QY 1198 GCCCTCTCTGCTTTGGCTCCATGAACGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1257
Db 360 ValIleGlyMetPheThrAlaLeuIleGlyMetMetLeuAlaGlySerArgLeuLeuTyr 379
QY 1258 GTTGGCTCTGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCCAAGCAC 1317
Db 380 SerPheGlyArgAspGlyLeuLeuProSerTrpLeuSerGlnLeuAsn---HisLysHis 398
QY 1318 ACTCCTCTACCAGCTGTTATTGTTTGTGCACCCCTTTGACAAATGATAATGCTCTCTCTGA 1377
Db 399 LeuProAsnArgAlaLeuAlaIle-----LeuThrIleIleGlyValValIleGly 415
QY 1378 GAC-----CTCGACAGCTCTTTTGAATTTCTCCTCAGTTTGGCAGGTGG 1419
Db 416 SerMetPheProPheAlaPheLeuAlaGlnLeuIleSerAlaGlyThrLeuValAlaPhe 435
QY 1420 CTTTTTATTGGGCTGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATG 1479

Db 436 MetPheValSerLeuAlaMetTyrArgLeu-----ArgLysArgGluGlyLysAspLeu 453
QY 1480 CATGCTCCT---TTCAAGGTGCCACTGTTT-----ATCCCAGCTTTGTTTTCCTTCACA 1530
Db 454 ProLysProGluPheLysLeuProLeuTyrProLleLeuProAlaIle---ThrPheIle 472
QY 1531 TGCCTCTTCATGTTGCCCTTTCCCTC---TATTCGGACCCATTAGTACAGGATTGGC 1587
Db 473 LeuValLeuValPheTrpGlyLeuSerPheGluAlaLysLeuTyrThrLeuIleTrp 492
QY 1588 TCGTCTCATCTGACTGGAGTCCCTGCTATCTCTTATCTCTTATATATATGGGACAGAAA 1647
Db 493 PheIleVal-----GlyIleIleIleTyrLeuIleTyrGlyIleArgHisSerLys 509

RESULT 11
US-08-132-990A-4
; Sequence 4, Application US/08132990A
; Patent No. 5834589
; GENERAL INFORMATION:
; APPLICANT: MERUELO, DANIEL
; APPLICANT: YOSHIMOTO, TAKAYUKI
; TITLE OF INVENTION: Human Retrovirus Receptor and DNA Coding Therefor
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/132,990A
; FILING DATE: 07-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/084,729
; FILING DATE: 29-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05569
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/899,075
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/806,178
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/627,950
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8105-004-999
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-132-990A-4

Alignment Scores: 1.41e-20 Length: 622
Pred. No.:

Score: 279.50 Matches: 102
Percent Similarity: 42.82% Conservative: 80
Best Local Similarity: 24.00% Mismatches: 176
Query Match: 6.78% Indels: 67
DB: 2 Gaps: 13
US-09-667-170A-440 (1-2239) x US-08-132-990A-4 (1-622)
QY 331 AGAGAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCAT 390
Db 23 ArgGluGluSerArgLeuSerArgCysLeuAsnThrTyrAspLeuValAlaLeuGlyVal 42
QY 391 GGCACCATCATTTGGAGCAGGAATCTTCATCTCTCTTAAGGGCGTCTCCAGAACACGGGC 450
Db 43 GlySerThrLeuGlyAlaGlyValTyrValLeuAlaGlyAlaValAlaArgGlu----- 60
QY 451 AGCGTGGGCATGTCTCTGACCATC---TGGACGGTGTGTGGGGTCTCTGTCACATTTTGA 507
Db 61 AsnAlaGlyProAlaIleValIleSerPheLeuIleAlaAlaLeuAlaSerValLeuAla 80
QY 508 GCTTGTCTTATGCTGAATTGGGAACAACACTATAAGAAATCTGGAGGTCATTACACATAT 567
Db 81 GlyLeuCysTyrGlyGluPheGlyAlaArgValProLysThrGlySerAlaTyrLeuTyr 100
QY 568 ATTTTGGAGTCTTTGCTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTC--- 624
Db 101 SerTyrValThrValGlyGluLeuTrpAlaPheIleThrGlyTrpAsnLeuIleLeuSer 120
QY 625 ---ATAATACGCCCTGCAGCTACTGCT---GTGATATCCCTGGCATTTGGACGCTACATT 678
Db 121 TyrIleIleGlyThrSerSerValAlaArgAlaTrpSerAlaThrPheAspGluLeuIle 140
QY 679 CTGGAACCATTT-----TATTTATCAA 699
Db 141 GlyLysProIleGlyGluPheSerArgGlnHisMetAlaLeuAsnAlaProGlyValLeu 160
QY 700 TGTGAAATCCCTGAACCTTGGATCGAGTCCCGGAGCGCCGATCCAGATTTTCTTAACCTTTGC 819
Db 177 GlyLeuLeuThrLeuGlyValLysGluSerAlaMetValAsnLysIlePheThrCysIle 196
QY 820 AAGCTCACAGCAATCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAA 879
Db 197 AsnValLeuValLeuCysPheIleValValSerGly-----PheValLysGlySer 213
QY 880 ACGCAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTG 939
Db 214 IleLysAsnTrpGln-----LeuThrGluLysAsnPheSerCysAsnAsnAspThr 231
QY 940 GCTTTTATTATGGA----- 954
Db 232 AsnValLysTyrGlyGluGlyGlyPheMetProPheGlyPheSerGlyValLeuSerGly 251
QY 955 -----ATGTATGCATATGCTGGCTGGTGTTCCTCACTCACTTGTACTGAAGAA 1002
Db 252 AlaAlaThrCysPheTyrAlaPheValGlyPheAspCysIleAlaThrThrGlyGluGlu 271
QY 1003 GTAGAAAACCTGAAAAAACCATTTCCCTTGCATATATGTATATCCATGGCCATTGTACC 1062
Db 272 ValLysAsnProGlnLysAlaIleProValGlyIleValAlaSerLeuLeuIleCysPhe 291
QY 1063 ATTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTG 1122
Db 292 Ile-----AlaTyrPhe---GlyValSerAlaAlaLeuThrLeu 303
QY 1123 CTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC----- 1173
Db 304 MetMetProTyrPheCysLeuAspIleAspSerProLeuProGlyAlaPheLysHisGln 323
QY 1174 -----TCATTAGCAGTTCCGATCTTGTGCCCCCTCTCCTGCTTT 1212

Db 324 GlyTrpGluGluAlaLysTyrAlaValAlaIleGlySerLeuCysAlaLeuSerThr--- 342
QY 1213 GGTCCCATGAACGGTGTGTTTGGCTGCTCCAGGTTATTCTATGTTGGCTCTCGAGAG 1272
Db 343 ---SerLeuLeuGlySerMetPheProMetProArgValIleTyrAlaMetAlaGluAsp 361
QY 1273 GGTACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCAGACTCTCTACACGCT 1332
Db 362 GlyLeuLeuPheLysPheLeuAlaLysIleAsnArgThrLysThrProValIleAla 381
QY 1333 GTTATTGTTTGCACCCCTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTT 1392
Db 382 ThrValThrSerGlyAlaIleAlaAlaValMetAlaPheLeuPheGluLeuLysAspLeu 401
QY 1393 TTGAATTCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTGTGCTGGGCTGATT 1452
Db 402 ValAspLeuMetSerIleGlyThrLeuLeuAlaTyrSerLeuValAlaAlaCysValLeu 421
QY 1453 TATCTTCGATACAAA 1467
Db 422 ValLeuArgTyrGln 426

RESULT 12
PCT-US92-09382-4
; Sequence 4, Application PC/TUS9209382
; GENERAL INFORMATION:
; APPLICANT: MERUELO, DANIEL
; APPLICANT: YOSHIMOTO, TAKAYUKI
; TITLE OF INVENTION: Human Retrovirus Receptor and DNA Coding
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09382
; FILING DATE: 19921213
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Livnat, Shmuel
; REGISTRATION NUMBER: 33,949
; REFERENCE/DOCKET NUMBER: MERUELO-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 628-5197
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-09382-4

Alignment Scores:
Pred. No.: 1,41e-20 Length: 622
Score: 279.50 Matches: 102
Percent Similarity: 42.82% Conservative: 80
Best Local Similarity: 24.00% Mismatches: 176
Query Match: 6.78% Indels: 67
DB: 5 Gaps: 13

US-09-667-170A-440 (1-2239) x PCT-US92-09382-4 (1-622)

QY 331 AGAGAGAAAGTGCAGCTCAAGAGAGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCAT 390
Db 23 ArgGluGluSerArgLeuSerArgCysLeuAsnThrTyrAspLeuValAlaLeuGlyVal 42
QY 391 GGCACCATCATTTGGAGCAGGAATCTTCTATCTCTCTCTAAGGGCGTCTCCAGAACACGGGC 450
Db 43 GlySerThrLeuGlyAlaGlyValTyrValLeuAlaGlyAlaValAlaArgGlu----- 60
QY 451 AGCGTGGGCATGTCTCTGACCATC--TGGACGGGTFGTGGGTCTCTGTCACATATTTGGA 507
Db 61 AsnAlaGlyProAlaIleValIleSerPheLeuIleAlaAlaLeuAlaSerValLeuAla 80
QY 508 GCTTTGCTTATGCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCAATACACATAT 567
Db 81 GlyLeuCysTyrGlyGluPheGlyAlaArgValProLysThrGlySerAlaTyrLeuTyr 100
QY 568 ATTTTGGAAAGTCTTGGTCCATTACCATCTTTGTACGAGTCTGGGTGGAACTCCTC--- 624
Db 101 SerTyrValThrValGlyGluLeuTrpAlaPheIleThrGlyTrpAsnLeuIleLeuSer 120
QY 625 ---ATAATACGCCCTGCAGCTACTGCT---GTGATATCCCTGGCATTTTGGACGCTACATT 678
Db 121 TyrIleIleGlyThrSerSerValAlaArgAlaTrpSerAlaThrPheAspGluLeuIle 140
QY 679 CTGGAACCATTT-----TTTATTTCAA 699
Db 141 GlyLysProIleGlyGluPheSerArgGlnHisMetAlaLeuAsnAlaProGlyValLeu 160
QY 700 TGTGAAATCCCTGAACCTTGGATCAAGCTCATACAGCTGTGGGCATAAAGTGTGATG 759
Db 161 AlaGlnThrProAspIlePheAlaValIleIle-----IleIleIleLeuThr 176
QY 760 GTCTAAATAGCATAGTGTGAGTGGAGCGCGCGGATCCAGATTTTCTTAACCTTTTGC 819
Db 177 GlyLeuLeuThrLeuGlyValLysGluSerAlaMetValAsnLysIlePheThrCysIle 196
QY 820 AAGCTCACAGCAATTCGATAAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAA 879
Db 197 AsnValLeuValLeuCysPheIleValValSerGly-----PheValLysGlySer 213
QY 880 ACGCAGAACTTTAAAGACGCGCTTTTCAGGAAGAGATTCAGATTTACCGCGGTGCCACTG 939
Db 214 IleLysAsnTrpGln-----LeuThrGluLysAsnPheSerCysAsnAsnAspThr 231
QY 940 GCTTTTATTATGGA----- 954
Db 232 AsnValLysTyrGlyGluGlyGlyPheMetPropheGlyPheSerGlyValLeuSerGly 251
QY 955 -----ATGTATGCATATGCTGGCTGGTGGTTTACCTCAACTTTGTACTGAAGAA 1002
Db 252 AlaAlaThrCysPheTyrAlaPheValGlyPheAspCysIleAlaThrThrGlyGluGlu 271
QY 1003 GTAGAAACCCCTGAAAAAACCATTCCTCTTGCATATATATATATATATATATATATATAT 1062
Db 272 ValLysAsnProGlnLysAlaIleProValGlyIleValAlaSerLeuLeuIleCysPhe 291
QY 1063 ATTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCAATTAATGTGAGGAGCTGCTG 1122
Db 292 Ile-----AlaTyrPhe---GlyValSerAlaAlaLeuThrLeu 303
QY 1123 CTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTTC----- 1173
Db 304 MetMetProTyrPheCysLeuAspIleAspSerProLeuProGlyAlaPheLysHisGln 323
QY 1174 -----TCATTAGCAGTTCGATCTTTGTGTCCTCTCTCTCTCTCTCTCTCTCTCT 1212
Db 324 GlyTrpGluGluAlaLysTyrAlaValAlaIleGlySerLeuCysAlaLeuSerThr--- 342
QY 1213 GGTCCCATGAACGGTGTGTTTGGCTGCTCCAGGTTATTCTATGTTGGCTCTCGAGAG 1272
Db 343 ---SerLeuLeuGlySerMetPheProMetProArgValIleTyrAlaMetAlaGluAsp 361
QY 1273 GGTACCTTCCAGAAATCT 1332

Db 362 GlyLeuLeuPheLysPheLeuAlaLysIleAsnAsnArgThrLysThrProValIleAla 381
 QY 1333 GTTATTGTTTGCACCCCTTTGACAATGATAATGCTCTCTCTCTGGAGACCTCGACAGTCTT 1392
 Db 382 ThrValThrSerGlyAlaIleAlaAlaValMetAlaPheLeuPheGluLeuLysAspLeu 401
 QY 1393 TTGAATTCCTCAGTTTGGCAGGTGGCTTTTATTTGGCTGGCAGTGTGGCTGATT 1452
 Db 402 ValAspLeuMetSerIleGlyThrLeuLeuAlaTyrSerLeuValAlaAlaCysValLeu 421
 QY 1453 TATCTTCGATACAAA 1467
 Db 422 ValLeuArgTyrGln 426

RESULT 13
 US-09-134-001C-5370
 ; Sequence 5370, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 5370
 ; LENGTH: 499
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-5370

Alignment Scores:
 Pred. No.: 7,68e-15 Length: 499
 Score: 224.50 Matches: 122
 Percent Similarity: 41.90% Conservative: 90
 Best Local Similarity: 24.11% Mismatches: 193
 Query Match: 5.45% Indels: 101
 DB: 4 Gaps: 26
 US-09-667-170A-440 (1-2239) x US-09-134-001C-5370 (1-499)
 QY 379 TCATTATCATTTGGCACCATTGAGCAGGAAATCTTCATCTCTCTAAGGCGTGTCTC 438
 Db 27 AlaIleAlaTyrGlySerCysIleGlyTrpGlyAlaPheIleLeuProGlyAspTrpIle 46
 QY 439 CAGAACACGGCAGCGTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGCTCTGTCA 498
 Db 47 LysGlnSerGlyProIleAlaSerSerIleGlyIle---ValIleGlyAlaLeuLeuMet 65
 QY 499 CTATTGGAGCTTGTCTTATGCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCTAT 558
 Db 66 IleLeuIleAlaValSerTyrGlyAlaLeuValGluLysPheProValSerGlyGlyAla 85
 QY 559 TACACATATATTTTGAAGTCTTTGGTCCATTACCATCTTTGTACAGTCTGGTGGAA 618
 Db 86 PheAlaPheSerPheLeuSerPheGlyArgTyrValSerPhePheSerSerTrpPhe--- 104
 QY 619 CTCCTCATATAACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATT 678
 Db 105 LeuThrPheGlyTyrValCysValValAlaLeuAsnAlaThrAlaPheSer----- 121
 QY 679 CTGGAACCATTTTTCATCAATGTGAATCCCTGAACTTGGCATTC-----AAGCTCATT 732
 Db 122 -----LeuLeuIleLysPheLeuLeuProAsnValLeuAsnGlyLysLeuTyr 138
 QY 733 ACAGCTGTGGC-----ATAACT 750
 Db 111

Db 139 ThrValAlaGlyTyrAspValTyrIleThrGluLeuValIleAlaThrValLeuLeu 158
 QY 751 GTAGTGATGGTCTCTAAATAGCATGAGTGTCTAGCTGGAGCGCCCGGATCCAGATTTCTTA 810
 Db 159 ValPheMetLeuIleThrIleArgGlyAlaSerValSerGlySerLeuGlnTyrTyr--- 177
 QY 811 ACCTTTTGAAGCTCACAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGCAGCTAATT 870
 Db 178 ---PheCys-----ValAlaMetValLeuValAlaLeuMet----- 189
 QY 871 AAAGGTCAAACGCGAAGCTTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGCGG 930
 Db 190 -----PheIleGlySerPhePheSerHisPheSerLeuSerHis 203
 QY 931 TTG---CCACTGGCT-----TTTATTATTGGATGTATGCA 963
 Db 204 LeuGluProLeuAlaSerValAspLysGlyTrpPheGlnSerIleIleMetIleValSer 223
 QY 964 TATGCTGGCTGGTCTTACCTCAACTTT-----GTTACTGAAGAAGTAGAA 1008
 Db 224 IleAlaProTrpAlaTyrValGlyPheAspAsnIleProGlnThrAlaGluGluPheAsn 243
 QY 1009 ---AACCTTGAAGAAACCATTCCTCTGCAATATATATATCCATGGCCATGTCCACCAT 1065
 Db 244 PheSerProAsnLysThrPheLysLeu---IleValTyrSerLeuAlaAlaSerLeu 262
 QY 1066 GGCTATGTGCTGACAAATGTG-----GCCTACTTTACGACCATTAAT 1107
 Db 263 ThrTyrValValMetLeuLeuTyrThrGlyTrpLeuSerThrGlnAlaThrSerLeuAsn 282
 QY 1108 GCTGAGGAGCTGCTGCTTCAATGCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGA 1167
 Db 283 Gly---AsnLeuTrpLeuThrGlyAlaVal-----ThrGlnAspAlaPheGly 297
 QY 1168 AATTCTCTATTAGCAGTT---CCGATCTTTGTTGCTCTCTCTCTGCTTTGGCTCCATGAAC 1224
 Db 298 PheIleGlyLeuAlaValLeuAlaValAlaIleIleMetGlyIlePheThrGlyLeuAsn 317
 QY 1225 GGTGGTGTGTTGCTGCTCCAGGTATTCTATGTTGCGTCTCGAGAGGGTCCACCTTCCA 1284
 Db 318 GlyPheLeuMetSerSerSerArgLeuLeuPheSerMetGlyArgSerGlyIleMetPro 337
 QY 1285 GAAATCCTCTCCATGATTCTGTCGCGAAGCACACTCTCTACCATGCTGTTATTTGTTG 1344
 Db 338 ThrValPheSerLysLeuHisSerLysHisLysThrProTyrValAlaIleIlePheLeu 357
 QY 1345 CACCCCTTTGACAATGATAATGCTCTCTCTCTGGA-----GAC 1380
 Db 358 ValAlaValSerLeuIleAlaProTrpLeuGlyArgThrAlaLeuThrTrpIleValAsp 377
 QY 1381 CTCGACAGTCTTTTGAATTTCTCTCAGTTTGGCAGGTGGCTTTTATTTGGCTGGCAGTT 1440
 Db 378 MetSerSerThrGlyValSerIleAlaTyr-----PheIleThrCysLeuSer 393
 QY 1441 GCTGGGCTGATTTATCTTCGATACAAATGCCACAGATATGCATCGTCT---TTCAAGGTG 1497
 Db 394 AlaThrLysLeuPheSerPheAsnLysGlnSerAsnThrTyrAlaProValTyrLysIle 413
 QY 1498 CCAGTGTTCATCCAGCTTTG-----TTTCTCTTCACATGCCTCTTCATGTT----- 1545
 Db 414 PheGlyIleIleGlySerIleValSerPheValPheLeuCysLeuLeuIleProGly 433
 QY 1546 -----GCCCTTTCCCTCTATTTCGGACCCCATTTTAGTACAGGATTTGGCTTCATC 1596
 Db 434 SerProAlaAlaLeuSerIleProSer---TyrIleAlaLeuGlyIleTrpLeuValIle 452
 QY 1597 ACTCTGACTGGAGTCCCTGCGTATTATCTCTTTTATATATGGACAAAGAACCCAGGTGG 1656
 Db 453 GlyLeu-----IlePhePheIleIle-----ArgLeuProLysLeu 464
 QY 1657 TTAGAATAATGTCAGAGAAATAACAGAACATTA-----CAAATAATA 1701
 Db 465 LysLysMetAsnAsnAspGluLeuSerArgLeuIleLeuAsnHisSerGluAspGluVal 484

Thu Apr 17 09:48:24 2003

QY 1702 CTGGAAGTGTACAGAA 1719
Db 485 LeuGluMetValHisGlu 490

RESULT 14
US-09-134-001C-4475
; Sequence 4475, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4475
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4475

Alignment Scores: 2.56e-12 Length: 494
Pred. No.: 97 Matches: 97
Score: 200.50
Percent Similarity: 43.02% Conservative: 88
Best Local Similarity: 22.56% Mismatches: 176
Query Match: 4.87% Indels: 69
DB: 4 Gaps: 20

US-09-667-170A-440 (1-2239) x US-09-134-001C-4475 (1-494)

QY 547 TCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACCAGCTTTTGTACGA 606
Db 8 SerGlyGlyMetSerAsnTyrAlaGlnTyrThrHisGlySerLeuLeuGlyPheIleAla 27
QY 607 GTC-----TGGGTGGAACCTCCTCATATACGCCCT-----GCAGCTACTGCTGTG 651
Db 28 AlaTyrAlaAsnTyrValSerLeuValThrIleIleProIleGluAlaValSerAlaVal 47
QY 652 ATATCCCTGGCATTTGGACGCTACATT-----CTGGAACCATTT 690
Db 48 -----GlnTyrMetSerSerTyrProTyrAspTyrAlaLysPro--- 60
QY 691 TTTATTCAATGTAATCCCTGAATCTGCGATCAAGCTCATTACAGCTGTGGGCATAACT 750
Db 61 ---MetGlySerLeuMetLysAspGlySerIleSerThrTyrGlyLeuIleAlaValTyr 79
QY 751 GTAGTGATGGTCTTAAATAGCATAGTGTGAGCTGGAGCGCCCGGATCCAGATTTCTTA 810
Db 80 IleIleIleAlaIlePheSerLeuLeuAsnTyrTyrSer-----ValLysLeuLeuThr 97
QY 811 ACCTTTGGCAAGCTCAGCAATCTGTGATAATTATAGTCCCT-----GGAGTTATG 861
Db 98 SerPheThrSerLeuIleSerValPheLysLeuGlyValProIleLeuThrIleIleMet 117
QY 862 CAGCTAATAAAGGT---CAAAACGAGCAACTTAAAGACGCCCTTTTCAGGAAGAGATTCA 918
Db 118 LeuLeuValSerGlyPheAspThrGlyAsnTyrGlyHisSerIleGly----- 133
QY 919 AGTATTACGGGTGGCCACTGGCTTTTATTATGGA----- 954
Db 134 -----ThrPheMetPro-----TyrGlySerAlaProIlePheAlaAlaThr 147
QY 955 -----ATGTATGCATATGCTGGCTGGTTTACCTCAACTTGTACTGAA 999
Db 148 ThrThrSerGlyIleIlePheSerPheAsnAlaPheGlnThrIleIleAsnMetGlySer 167

QY 1000 GAAGTAGAAACCCCTGAAAAAACCATTCCTCCCTTGAATATGTATATATCCATGGCCATTGTC 1059
Db 168 GluIleLysAsnProGluLysAsnIleAlaArgGlyIleValIleSerLeuThrLeuSer 187
QY 1060 ACCATTGGCTATGTGCTGACAAAATGTGGCTACTTTTACGACCAATTAATGCTGAGGAGCTG 1119
Db 188 AlaIleLeuTyrIleValLeuGlnSerThrPheIleThrSerMetProSerSerMetLeu 207
QY 1120 CTGCTTTCAAATGCAGTGGCAGTGACCTTCTTCTGAGCGG-----CTA 1161
Db 208 HisGluHisGlyTyrSerGlyIleAsnPheAsnSerPropheAlaAspMetAlaIleLeu 227
QY 1162 CTGGGAATTTCTCATTAGCAGTTCCGATCTTTGTT-----GCCCTCTCTGCTTTGGC 1215
Db 228 LeuGlyLeuAsnTyrLeuAlaIleLeuLeuTyrMetGluAlaValValSerPropheGly 247
QY 1216 TCCATGAACGGTGGTGTGTTGCTGTCTCCAGTTTATCTATGTTGCGCTCGAGAGGGT 1275
Db 248 ThrGlyValSerPheValAlaValThrGlyArgValLeuArgAlaMetGluGluAsnGly 267
QY 1276 CACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACCTCTCTACAGCTGTT 1335
Db 268 HisIleProLysPheLeuGlyLysIleAsnLysLysTyrAsnIleProArgValAlaIle 287
QY 1336 ATT-----GTTTTCACCCCTTTGACAAATGATAATGCTCTTCTGGAGACCTCGACAGT 1389
Db 288 AlaPheAsnAlaIleIleSerMetValMetValThrLeuPheArg---AspTyrGlyThr 306
QY 1390 CTTTGAATTTCTCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTG 1449
Db 307 LeuAlaAlaValIleSerThrAlaThrLeuValAlaTyrLeuThrGlyProThrVal 326
QY 1450 ATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGGTGCCACTGTTT 1506
Db 327 IleSerLeuArgLysMetAlaProLysMetThrArgPropheLysAlaAsnIleLeuLys 346
QY 1507 ATCCCAGCTTTGTTTCTCCTCACATGCTCTTCAATGTTGCTTCCCTTCTTATTCGGAC 1566
Db 347 PheMetAlaProLeuSerPheVal-----LeuAlaSerLeuAlaIleTyrTyrAla 363
QY 1567 CCATTT---AGTACAGGAGTGGCTTCGCTCATCTCATCTGAGTGGAGTCCCTGCGTATTAT 1623
Db 364 MetTyrProThrThrAlaGluValIleIleIleIleIleLeuGlyLeuProIleTyrPhe 383
QY 1624 CTCTTTTATATATGGACAAAGAACCCAGGTGGTGTAGATAATGTCAGAGAAATAAAC 1683
Db 384 PheTyr-----GluTyrLysMetAsnTyr---LysAsnThrLysLysGlnIleGly 399
QY 1684 AGAACATTACAAATAATACTGGAAGTTGTA 1713
Db 400 GlySerLeuTyrIleIleIleIleTyrLeuIle 409

RESULT 15
US-09-134-001C-3778
; Sequence 3778, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3778
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3778

Alignment Scores:

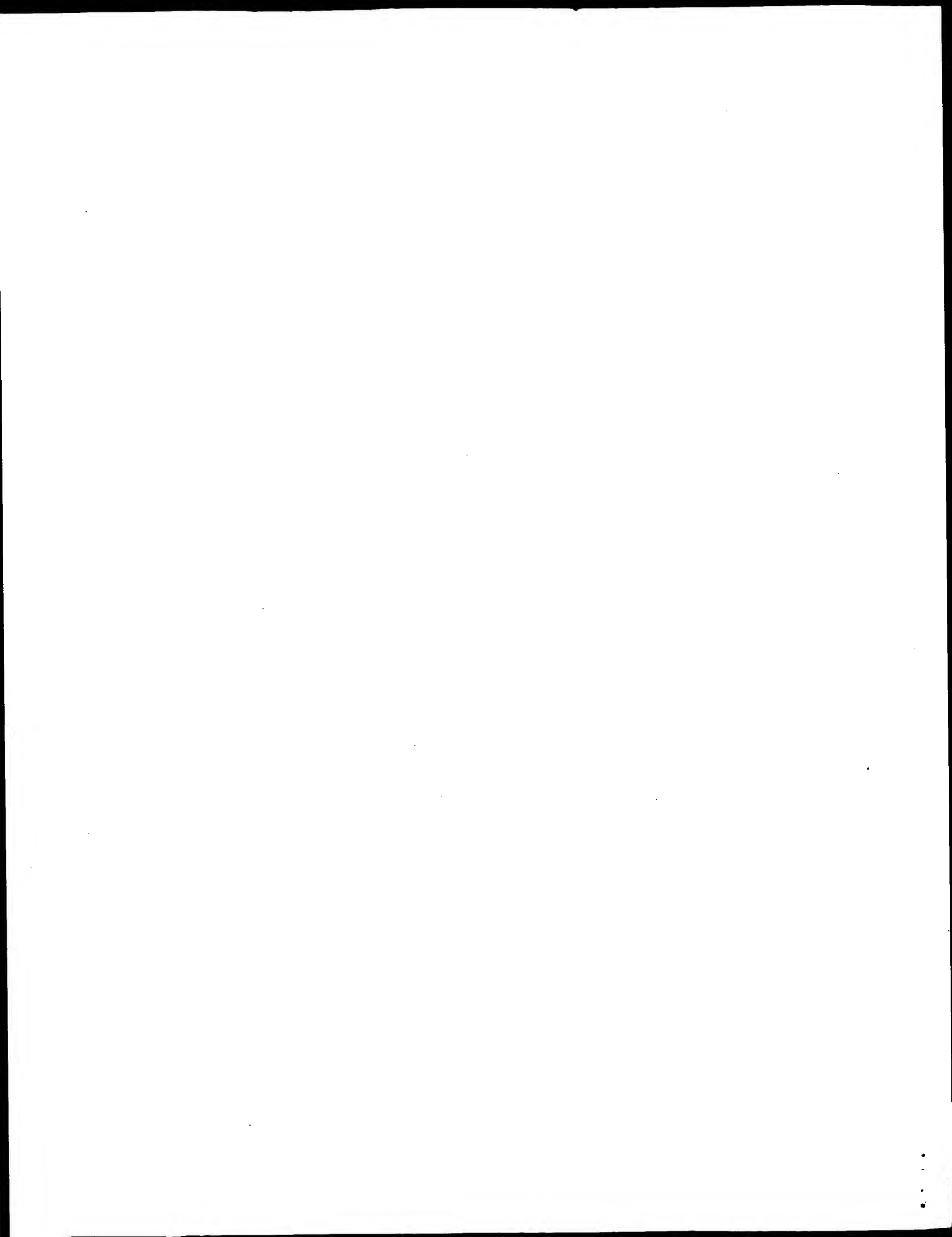
Pred. No.: 7.48e-12 Length: 476
Score: 196.00 Matches: 100
Percent Similarity: 38.71% Conservative: 80
Best Local Similarity: 21.51% Mismatches: 213
Query Match: 4.76% Indels: 72
DB: 4 Gaps: 18

US-09-667-170A-440 (1-2239) x US-09-134-001C-3778 (1-476)

QY 370 AGGGGAGTCTCCATTATC---ATTGGCACCATCATATTGGAGCAGGAATCTTCATCTCTCCT 426
||| :|||:|||| ||||| ||||| |||||:||||:||||: |||||
Db 33 ArgHisIleSerMetIleAlaIleGlyGlyCysIleGlyThrGlyLeuPheMetThrSer 52
QY 427 AAGGCGGTGCTCCAGAACACGGGACGGTGGGATGCTCTCTGACCATCTGGACGGTGTGT 486
||| :|||:||||:||||: ||||| ||||| |||||:||||: |||||
Db 53 GlyGlyAlaIleHisAspAlaGlyAlaLeuGly---AlaLeuIleAlaTyrAlaValIle 71
QY 487 GGGGTCTGTCACTATTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAAAGAAA 546
||| :|||:|||| ||||| ||||| |||||:||||: |||||
Db 72 GlyAlaMetValPhePheLeuMetThrSerLeuGlyGluMetAlaThrTyrLeuProVal 91
QY 547 TCTGGAGGTCAATTACACATATATTTTGGAACTCTTTGGTCCATTACCAGCTTTTGTACGA 606
||||| ||||| ||||| ||||| ||||| |||||:||||: |||||
Db 92 SerGlySerPheSerThrTyrAlaThrArgPheValAspProSerLeuGlyPheAlaLeu 111
QY 607 GTCTGGGTGGAACCTCTC-----ATAATACGCCCTGCAGCTACTGCTGTGATATCCCTG 660
||| :|||:|||| ||||| ||||| |||||:||||: |||||
Db 112 GlyTyrAsnTyrTrpPheAsnTrpValIleThrValAlaAlaAspValThrIleAlaAla 131
QY 661 GCATTTGGACGCTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACTTGGC 720
||| :|||:|||| ||||| ||||| |||||:||||: |||||
Db 132 Gln-----ValIleGlnTyrTrpSerProMetGlnGly 142
QY 721 ATCAAG-----CTCATTACAGCTGTGGGCATAACTGTAGTGTGCTCTAAATAGCATG 774
||| :|||:|||| ||||| ||||| |||||:||||: |||||
Db 143 IleProAlaTrpValTrpSerCysIlePheLeuIleIleIlePheAlaLeuAsnSerLeu 162
QY 775 AGTGTACAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGCAGCTCACAGCAATT 834
||||| ||||| ||||| ||||| |||||:||||: |||||
Db 163 SerValArgValTyrGlyGluSerGluTyrTrpPheAlaLeuIleLysValValThrVal 182
QY 835 CTGATAATTATAGTCCCTGGAGTTATGCAGCTAATT-----AAAGGTCAAACGCGAG 885
:|||| ||||| |||||:||||: ||||| |||||:||||: |||||
Db 183 IleIlePheIleGlyIleGlyIleLeuThrIleLeuGlyIleMetGlyGlyGluPheVal 202
QY 886 AACTTTAAAGACGCCCTTTTCAGGAACAGAT-----TCA 918
||| ||| |||||:||||: |||||
Db 203 GlyPhe---AspThrPheThrLysGlyAspGlyProIleLeuGlyGlyAsnLeuGlyGly 221
QY 919 AGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGATGATATGCTGGCTGGTTT 978
||||: ||||| ||||| |||||:||||: |||||
Db 222 SerLeuLeuSerIleLeuGlyValPheLeuValAlaGlyPheSerPheGlnGlyThrGlu 241
QY 979 TACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAAACCATTTCCCTTGCATA 1038
:|||| ||||| |||||:||||: ||||| |||||:||||: |||||
Db 242 LeuIleGlyIleThrAlaGlyGluSerGluAsnProGluArgAlaValProLysAlaIle 261
QY 1039 TGTATATCCATGGCCATTGTCCACCATTTGGCTATGCTGTGACAAATGTGGCCTACTTACG 1098
:|||| ||||| |||||:||||: |||||
Db 262 LysGlnValPheTrpArgIleLeuLeuPheTyrIleLeuAlaIlePheIleIleGlyMet 281
QY 1099 ACCATTAAATGCTGAGGAGCTGCTGCTT-----TCAAATGCAGTGGCAGTGACCTTT 1149
||| :|||:|||| ||||| |||||:||||: |||||
Db 282 LeuIleProTyrAspSerLysAlaLeuMetGlyGlyGlyAspSerIleAlaThrSerPro 301
QY 1150 TCTGAGCGGCTACTGGGAATTTCTCATTAGCA-----GTTCCGATCTTT 1194
:|||| ||||| |||||:||||: |||||
Db 302 PheThrLeuValPheLysAsnAlaGlyLeuAlaPheAlaAlaSerPheMetAsnAlaVal 321
QY 1195 GTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAGGTTATTC 1254
:|||| ||||| |||||:||||: |||||

Db 322 IleLeuThrSerValLeuSerAlaGlyAsnSerGlyMetTyrAlaSerThrArgMetLeu 341
QY 1255 TATGTTGGCTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAG 1314
||| |||||:||||: ||||| |||||:||||: |||||
Db 342 TyrSerMetSerLysAspLysLeuAlaTyrAsnSerPheGlyLysThrAsn---LysSer 360
QY 1315 CACACTCTCTACACAGCTGTTATTGTTTTCACCCCTTTTGACAAATGATAATGCTC----- 1368
||| |||||:||||: ||||| |||||:||||: |||||
Db 361 GlyValProTyrValSerLeuIleAlaThrGlyValLeuValIleLeuIlePheAlaLeu 380
QY 1369 -----TTCTCTGGAGAC-----CTCGACAGTCTTTTGAATTTCTCTCAGTTT 1410
||| ||||| ||||| |||||:||||: |||||
Db 381 GlnHisLeuSerGlyAspAlaTyrGluTyrIleValAlaAlaSerGlyMetThrGlyPhe 400
QY 1411 GCCAGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTAT----- 1455
||| |||||:||||: ||||| |||||:||||: |||||
Db 401 IleAlaTrp-----ValGlyIleAlaIleSerHisPheArgPheArgAlaPheAsp 418
QY 1456 -----CTTCGATACAAATGCCCAGATATGCATCGTCTCTTC 1491
||| ||||| ||||| |||||:||||: |||||
Db 419 LysGlnAsnTyrAspLysSerLysLeuLysTyrLysAla----- 431
QY 1492 AAGGTGCCACTGTTTCATCCAGCTTTGTTTTCCTTCACATGCCTCTTCATGGTGGCCTT 1551
||||: ||||| ||||| |||||:||||: |||||
Db 432 LysLeuPheProPheGlyProIlePheAlaGlyIleLeuCysVal---IleValIleIle 450
QY 1552 TCCCTCTATTCCGACCCCATTTAGTACAGGG-----ATTGGCTTCGTCATCACT 1599
||| ||||| ||||| |||||:||||: |||||
Db 451 GlyGlnAspValAspPheIleLysThrGlyAsnPheAspMetAsnArgPhePheIleThr 470
QY 1600 CTGACTGGAGTCCCT 1614
||| ||||| |||||
Db 471 SerMetGlyIlePro 475

Search completed: April 16, 2003, 16:27:03
Job time : 35 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: April 16, 2003, 16:10:40 ; Search time 65.5 Seconds
(without alignments)
5169.450 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 288829 seqs, 75613885 residues
Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_090919_20390/app_query.fasta_1.2375
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09667170 @cgn_1_1_15 @runat_04042003_090919_20390
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2519	61.1	501	9 US-10-163-866-48	Sequence 48, Appl
2	2519	61.1	501	9 US-10-163-866-49	Sequence 49, Appl
3	1210.5	29.4	507	9 US-10-163-866-54	Sequence 54, Appl
4	1206.5	29.3	524	9 US-10-163-866-37	Sequence 37, Appl

5	1198.5	29.1	507	9 US-10-163-866-38	Sequence 38, Appl
6	1198.5	29.1	507	9 US-10-163-866-39	Sequence 39, Appl
7	1156.5	28.1	515	9 US-10-163-866-40	Sequence 40, Appl
8	1131.5	27.5	511	9 US-10-163-866-41	Sequence 41, Appl
9	1131.5	27.5	511	9 US-10-163-866-42	Sequence 42, Appl
10	1101.5	26.7	535	9 US-10-163-866-43	Sequence 43, Appl
11	1101.5	26.7	535	9 US-10-163-866-44	Sequence 44, Appl
12	1075.5	26.1	487	9 US-10-163-866-46	Sequence 46, Appl
13	1067.5	25.9	487	9 US-10-163-866-45	Sequence 45, Appl
14	1048	25.4	517	9 US-09-815-923-16	Sequence 16, Appl
15	1039	25.2	523	9 US-10-163-866-47	Sequence 47, Appl
16	732	17.8	414	10 US-09-925-297-747	Sequence 747, App
17	513	12.5	97	9 US-09-854-133-586	Sequence 586, App
18	513	12.5	97	10 US-09-738-973-586	Sequence 586, App
19	412.5	10.0	180	9 US-10-163-866-50	Sequence 50, Appl
20	406.5	9.9	179	10 US-09-864-761-43216	Sequence 43216, A
21	322	7.8	456	10 US-09-815-242-4932	Sequence 4932, Ap
22	322	7.8	463	10 US-09-815-242-10662	Sequence 10662, A
23	295	7.2	619	10 US-09-741-153-4	Sequence 4, Appli
24	286.5	7.0	619	10 US-09-741-153-2	Sequence 2, Appli
25	262	6.4	470	10 US-09-815-242-12078	Sequence 12078, A
26	257.5	6.2	482	10 US-09-815-242-12941	Sequence 12941, A
27	233.5	5.7	463	10 US-09-815-242-13918	Sequence 13918, A
28	232	5.6	475	10 US-09-815-242-10052	Sequence 10052, A
29	230.5	5.6	453	9 US-09-738-626-4026	Sequence 4026, Ap
30	229	5.6	468	9 US-09-738-626-4777	Sequence 4777, Ap
31	225.5	5.5	149	9 US-10-043-487-376	Sequence 376, App
32	223.5	5.4	497	10 US-09-815-242-12626	Sequence 12626, A
33	219	5.3	469	10 US-09-815-242-5281	Sequence 5281, Ap
34	218.5	5.3	489	10 US-09-815-242-13932	Sequence 13932, A
35	216.5	5.3	489	10 US-09-815-242-10215	Sequence 10215, A
36	208.5	5.1	475	10 US-09-815-242-11503	Sequence 11503, A
37	207.5	5.0	519	9 US-09-895-913A-118	Sequence 118, App
38	207.5	5.0	519	10 US-09-815-242-11388	Sequence 11388, A
39	199.5	4.8	487	10 US-09-815-242-12036	Sequence 12036, A
40	197.5	4.8	459	10 US-09-815-242-12844	Sequence 12844, A
41	176	4.3	463	9 US-09-738-626-6946	Sequence 6946, Ap
42	174	4.2	603	10 US-09-801-368-144	Sequence 144, App
43	173	4.2	466	10 US-09-815-242-10265	Sequence 10265, A
44	173	4.2	475	10 US-09-815-242-5055	Sequence 5055, Ap
45	162	3.9	466	10 US-09-815-242-14048	Sequence 14048, A

ALIGNMENTS

RESULT 1
US-10-163-866-48
; Sequence 48, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-48

Alignment Scores:

Pred. NO.: 1.47e-237 Length: 501
 Score: 2519.00 Matches: 498
 Percent Similarity: 96.51% Conservative: 0
 Best Local Similarity: 96.51% Mismatches: 0
 Query Match: 61.14% Indels: 18
 DB: 9 Gaps: 1

US-09-667-170A-440 (1-2239) x US-10-163-866-48 (1-501)

QY	184	AAGCCTGTTGTTCCACCATCTCCAAAGGAGGTTACCTGAGGAAATGTTAACGGGAGG	243
Db	4	LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlyAsnValAsnGlyArg	23
QY	244	CTGCCTTCCCTGGGCAACAAGGAGGACACCTGGGACGAGCGCTTTTCAGGAAGAGACGCC	303
Db	24	LeuProSerLeuGlyAsnLysGluProGlyGln-----	35
QY	304	TTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCACT	363
Db	36	-----GluLysValGlnLeuLysArgLysValThr	45
QY	364	TTACTGAGGGGAGTCTCCATTATCATTTGGCACCACCATCATTTGGACGAGGAATCTTCATCTCT	423
Db	46	LeuLeuArgGlyValSerIleIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer	65
QY	424	CCTAAGGGCGTGTCTCCAGAACACGGGACGGCGTGGCATGTCTCTGACCATCTGGACGGTG	483
Db	66	ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal	85
QY	484	TGTGGGTCCTGTCTCACTATTGGAGCTTTTGGAGTCTTTGGTCCATTACCAAGTATAAAG	543
Db	86	CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrThrIleLys	105
QY	544	AAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGGTCCATTACCAAGTATAAAG	603
Db	106	LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal	125
QY	604	CGAGTCTGGTGGGACTCCTCATATACGACCTGACGCTACGCTGTGTGATATCCCTGGCA	663
Db	126	ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla	145
QY	664	TTTGGACGCTACATCTCTGGAAACCATTTTATCAATGTGAATCCCTGAACTTGGGATC	723
Db	146	PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle	165
QY	724	AAGCTCATTAAGCTGTGGGCATAACTGTAGTGTGATGCTCTAAATAGCATGAGTGTGAGC	783
Db	166	LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer	185
QY	784	TGGAGCGCGGATCCAGATTTTCTTAACCTTTTGCAGGTCACAGCAATTTCTGATAATT	843
Db	186	TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIle	205
QY	844	ATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGACGAGAACTTTAAAGACGCCCTT	903
Db	206	IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPhelLysAspAlaPhe	225
QY	904	TCAGGAAGAGATCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATGCA	963
Db	226	SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla	245
QY	964	TATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAACC	1023
Db	246	TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr	265
QY	1024	ATTCCTCTGCAATATGATATCATTCATGGCCATTTGTCACCATTTGGCTATGTGCTGACAAAT	1083
Db	266	IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn	285
QY	1084	GTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTG	1143
Db	286	ValAlaTyrPheThrThrIleAsnAlaGluGluLeuLeuSerAsnAlaValAlaVal	305

QY	1144	ACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCGGATCTTTGTTGCCCTC	1203
Db	306	ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu	325
QY	1204	TCCTGCTTTGGCTCCATGAACGGTGGTGTGTGTCTCCAGGTATCTATGTTGCG	1263
Db	326	SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla	345
QY	1264	TCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTCT	1323
Db	346	SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro	365
QY	1324	CTACCAAGCTGTTATGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACTC	1383
Db	366	LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu	385
QY	1384	GACAGTCTTTTGAATTTCTCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTGT	1443
Db	386	AspSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheIleGlyLeuAlaValAla	405
QY	1444	GGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCACTG	1503
Db	406	GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgPropheLysValProLeu	425
QY	1504	TTCATCCAGCTTTGTTTCTCTTCCATGCCTCTTTCATGGTGGCTTTTCCCTCTATTCTG	1563
Db	426	PheIleProAlaLeuPheSerPheThrCysLeuPheMetValAlaLeuSerLeuTyrSer	445
QY	1564	GACCAITTAGTACAGGATTTGGCTTCCATGCCTCTTTCATGGTGGCTGGAGTCCCTGCTATTAT	1623
Db	446	AspProPheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr	465
QY	1624	CTCTTTTATATATATGAGGATTTGGCTTCCATGCCTCTTTCATGGTGGCTGGAGTCCCTGCTATTAT	1683
Db	466	LeuPheIleIleTrpAspLysLysProArgTrpPheArgIleMetSerGluLysIleThr	485
QY	1684	AGAACATTACAAATAATCTGGAAGTTGTACCAGAAAGATAAGTTA	1731
Db	486	ArgThrLeuGlnIleIleLeuGluValValProGluGluAspLysLeu	501

RESULT 2

US-10-163-866-49
 ; Sequence 49, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: US 60/296,076
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 49
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-163-866-49

Alignment Scores:
 Pred. No.: 1.47e-237 Length: 501
 Score: 2519.00 Matches: 498
 Percent Similarity: 96.51% Conservative: 0

Best Local Similarity: 96.51% Mismatches: 0
Query Match: 61.14% Indels: 18
DB: 9 Gaps: 1
US-09-667-170A-440 (1-2239) x US-10-163-866-49 (1-501)
QY 184 AAGCTGTGTGTCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAAACGGGAGG 243
Db 4 LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlyAsnValAsnGlyArg 23
QY 244 CTGCCTTCCCTGGCAACAAGGAGCCACCTGGGCGAGGAGCCCTTTTCAGGAAGAGACGCC 303
Db 24 LeuProSerLeuGlyAsnLysGluProProGlyGln----- 35
QY 304 TTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACT 363
Db 36 -----GluLysValGlnLeuLysArgLysValThr 45
QY 364 TTACTGAGGGAGCTCTCCATTATCATTTGGCACCATCATTTGAGCAGGAAATCTTCATCTCT 423
Db 46 LeuLeuArgGlyValSerIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 65
QY 424 CCTAAGGGCGTCTCCAGAACACGGCGAGCGTGGGCGATGCTCTGACCATCTGGACGGTG 483
Db 66 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal 85
QY 484 TGTGGGCTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAG 543
Db 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrIleLys 105
QY 544 AAATCTGGAGGTCATTACACATATATTTTGAAGCTTTTGGTCCATTACCAGCTTTTGTA 603
Db 106 LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal 125
QY 604 CGAGTCTGGTGGAACTCCTCATATAATACGCCCTGAGCTACTGCTGTGATATCCCTGGCA 663
Db 126 ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla 145
QY 664 TTTGGACGCTACATCTGGAAACCATTTTATTAATGTGAATCCCTGAACCTTGGCATC 723
Db 146 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 165
QY 724 AAGCTCATACAGCTGTGGGCATAACTGTAGTGATGTCCTAAATACCATGAGTGTACGC 783
Db 166 LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 185
QY 784 TGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTTGAAGCTCAGCAATCTGATAAT 843
Db 186 TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIle 205
QY 844 ATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCGAGAACTTTAAAGACGCTTT 903
Db 206 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 225
QY 904 TCAGGAAGAGATTCAAGTATTACCGCGGTGGCCACTGGCTTTTATGGAATGTATGCA 963
Db 226 SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 245
QY 964 TATGCTGGCTGGTTTACCTCAACTTGTACTGAAGAGTGAAGAAACCCCTGAAAAAACCC 1023
Db 246 TyrAlaGlyTyrPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 265
QY 1024 ATTCCCTTGGCAATATGATATCATGCGCCATTGTACCATTTGGCTATGCTGACAAAT 1083
Db 266 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 285
QY 1084 GTGGCCTACTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTG 1143
Db 286 ValAlaTyrPheThrThrIleAsnAlaGluGluLeuLeuSerAsnAlaValAlaVal 305
QY 1144 ACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCCGATCTTTGTTGCCCTC 1203
Db 306 ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 325

QY 1204 TCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTCATGTTGCG 1263
Db 326 SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla 345
QY 1264 TCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTCCT 1323
Db 346 SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro 365
QY 1324 CTACCAGCTGTTATTTGTTTGGCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACCTC 1383
Db 366 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 385
QY 1384 GACAGTCTTTGAATTTCTCCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTTGCT 1443
Db 386 AspSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheIleGlyLeuAlaValAla 405
QY 1444 GGGCTGATTTATCTTCGATACAAAATGCCAGATATGCATCGTCTTTCAAGGTGCCACTG 1503
Db 406 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 425
QY 1504 TTTATCCCGAGCTTTGTTTCTTCCATGCATGCCTTTCATGGTTGCCCTTCCCTCTATTTCG 1563
Db 426 PheIleProAlaLeuPheSerPheThrCysLeuPheMetValAlaLeuSerLeuTyrSer 445
QY 1564 GACCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTAT 1623
Db 446 AspProPheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr 465
QY 1624 CTCTTTTATATATGGACAAAGAAACCCAGGTGGTGTAGAAATATGTCAGAGAAATAACC 1683
Db 466 LeuPheIleIleTrpAspLysLysProArgTrpPheArgIleMetSerGluLysIleThr 485
QY 1684 AGAACATTTACAAATAATACTGGAAGTTGTACCAAGAAAGATAAGTTA 1731
Db 486 ArgThrLeuGlnIleIleLeuGluValValProGluGluAspLysLeu 501

RESULT 3

US-10-163-866-54
; Sequence 54, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-54

Alignment Scores:
Pred. No.: 1,55e-109 Length: 507
Score: 1210.50 Matches: 232
Percent Similarity: 66.19% Conservative: 95
Best Local Similarity: 46.96% Mismatches: 164
Query Match: 29.38% Indels: 3
DB: 9 Gaps: 2

US-09-667-170A-440 (1-2239) x US-10-163-866-54 (1-507)

QY	247	CCTTCCCTGGGCAACAAGGAGCCACCTGGCAGGACGCTTTTTCAGGAAGACGCCTTT	306
Db	14	ProAlaAlaGluGluLysGluGluAlaArgGluLysMetLeuAlaAlaLysSerAlaAsp	33
QY	307	TCAGGAAGACACGCCTTTTCAGGAAGAGAGAAAGTCAGCTGAAGAGGAAAGTCACCTTA	366
Db	34	GlySerAlaProAlaGlyGluGly--GluGlyValThrLeuGlnArgAsnIleThrLeu	52
QY	367	CTGAGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCACTCTCTCCT	426
Db	53	LeuAsnGlyValAlaIleIleValGlyThrIleIleGlySerGlyIlePheValThrPro	72
QY	427	AAGGGCGTGTCTCCAGAACACGGCGAGCGTGGGCATGTCTCTGACCATCTGGACGGTGTGT	486
Db	73	ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValTrpAlaAlaCys	92
QY	487	GGGGTCTGTCTACTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATATAAGAAA	546
Db	93	GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleSerLys	112
QY	547	TCTGGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTGTACGA	606
Db	113	SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys	132
QY	607	CTCTGGGTGGAACCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTT	666
Db	133	LeuTrpIleGluLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe	152
QY	667	GGACGCTACATTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACTTGGCATCAAG	726
Db	153	AlaThrTyrLeuLeuLysProLeuPheProThrCysProValProGluAlaAlaLys	172
QY	727	CTCAATTACAGCTGTGGGCATAACTGTAGTATGGTCTTAATAGCATGAGTGTGAGCTGG	786
Db	173	LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla	192
QY	787	AGCGCCGGATCCAGATTTTCTTAACCTTTTGCAGCTCACAGCAATTCTGTATAATTATA	846
Db	193	AlaThrArgValGlnAspAlaPheAlaAlaAlaLysLeuAlaLeuAlaLeuIleIle	212
QY	847	GTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGAGAACTTTAAAGACGCTTTTCA	906
Db	213	LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer	232
QY	907	-----GGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTAT	960
Db	233	PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe	252
QY	961	GCATATGCTGGCTGGTTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCTGAAAAA	1020
Db	253	AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg	272
QY	1021	ACCATTCCCTTCGAATATGTATATCCATGGCCATTGTACACATTGGCTATGTGCTGACA	1080
Db	273	AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr	292
QY	1081	AATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCA	1140
Db	293	AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla	312
QY	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTTGTTGCC	1200
Db	313	ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly	332
QY	1201	CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAGGTTATTCTATGTT	1260
Db	333	LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPhePheVal	352
QY	1261	GGTCTCGAGAGGGTCACCTTCCAGAAATTCCTCATGATTTCATGCTCCGACACACT	1320
Db	353	GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr	372

QY	1321	CCTCTACCAAGCTGTTATTGTTTTTGCACCCTTTGACAATGATAATGCTCTTCTCTGGAGAC	1380
Db	373	ProValProSerLeuValPheThrCysValMetThrLeuLeuTyAlaPheSerLysAsp	392
QY	1381	CTCGACAGTCTTTTGAAATTCCTCAGTTTTGCCAGGTGGCTTTTATTATGGGCTGGCAGTT	1440
Db	393	IlePheSerValIleAsnPhePheSerPhePheAsnTrpLeuCysValAlaLeuAlaIle	412
QY	1441	GCTGGGCTGATTTATCTTCGATACAAAATGCCAGATFATGCATCGTCTTTCAAGGTGCCA	1500
Db	413	IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn	432
QY	1501	CTGTTCCATCCCAGCTTTTGTTTTCCTTCACATGCCTCTTCATGGTTGCCCTTCCCTCTAT	1560
Db	433	LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp	452
QY	1561	TCCGACCCCATTTAGTACAGGGATTTGGCTTCGTCATCACCTCTGACTGGAGTCCCTCGCTAT	1620
Db	453	LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr	472
QY	1621	TATCTCTTTATTATATGGGACAGAACCAGGTGGTTTAGAATAATGTACAGAGAAAAATA	1680
Db	473	PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr	492
QY	1681	ACCAGAACATTACAAAATAATACTGGAAGTTGTACCAGAAGAA	1722
Db	493	ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu	506
 RESULT 4 US-10-163-866-37 ; Sequence 37, Application US/10163866 ; Publication No. US20030027188A1 ; GENERAL INFORMATION: ; APPLICANT: EXELIXIS, INC. ; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF ; FILE REFERENCE: EX02-080C ; CURRENT APPLICATION NUMBER: US/10/163,866 ; CURRENT FILING DATE: 2002-06-05 ; PRIOR APPLICATION NUMBER: US 60/296,076 ; PRIOR FILING DATE: 2001-06-05 ; PRIOR APPLICATION NUMBER: US 60/328,605 ; PRIOR FILING DATE: 2001-10-10 ; PRIOR APPLICATION NUMBER: US 60/338,733 ; PRIOR FILING DATE: 2001-10-22 ; PRIOR APPLICATION NUMBER: US 60/357,253 ; PRIOR FILING DATE: 2002-02-15 ; PRIOR APPLICATION NUMBER: US 60/357,600 ; PRIOR FILING DATE: 2002-02-15 ; NUMBER OF SEQ ID NOS: 54 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 37 ; LENGTH: 524 ; TYPE: PRT ; ORGANISM: Homo sapiens US-10-163-866-37			

Alignment Scores:	
Pred. No.:	3.88e-109
Score:	1206.50
Percent Similarity:	66.19%
Best Local Similarity:	46.76%
Query Match:	29.28%
DR:	9
Length:	524
Matches:	231
Conservative:	96
Mismatches:	164
Indels:	3
Gaps:	2

PG-08-667-170A-440 (1-2239) X US-10-163-866-37 (1-524)

QY	247	CCTTCCCTGGCAACAAGGAGGCCACTGGGCAGGACGCCCTTTTCAGGAAGAGACGCCTTT	306
		:: :: :: :: :: ::	
Dβ	31	proAlaAalaglLysLysLysGluAlaArgGluLysMetLeuAlaAlaLysSerAlaAsp	50
QY	307	TCAGGAAGAGACGCCTTTTCAGGAAGACAGAAAGTGCGAGCTGAAGAGGAAAGTCACITTA	366

Db 51 GlySerAlaProAlaGlyGluGly---GluGlyValThrLeuGlnArgAsnIleThrLeu 69

QY 367 CTGAGGGGAGTCTCCATTATATATGCGACCATCATTTGGAGCAGGAATCTTCATCTCTCCT 426

Db 70 LeuAsnGlyValAlaIleValGlyThrIleIleGlySerGlyIlePheValThrPro 89

QY 427 AAGGGCGTGCTCCAGAACACGGGCGAGCGTGCGCATGTCTCTGACCATCTGGACGGTGTGT 486

Db 90 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValThrAlaAlaCys 109

QY 487 GGGGTCCTGTACATATTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAAAGAAA 546

Db 110 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrIleSerLys 129

QY 547 TCTGGAGGTCAATACACATATATTTTGGAGTCTTTGGTCCATTACCAAGCTTTTGTACGA 606

Db 130 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 149

QY 607 GTCTGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTT 666

Db 150 LeuTrpIleGluLeuLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe 169

QY 667 GGACGCTACATCTCTGGAACCATTTTATTTTCAATGTGAATCCCTGAACCTTGCATCAAG 726

Db 170 AlaThrTyrLeuLeuLysProLeuPheProThrCysProValProGluAlaAlaLys 189

QY 727 CTCATTACAGCTGTGGGCATAACTGTAGTGATGGTCTCTAAATAGCATGAGTGTGAGCTGG 786

Db 190 LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAlaAsnCysTyrSerValLysAla 209

QY 787 AGCGCCCGGATCCAGATTTTCTTAACCTTTTGCAAGCTCACAGCAATCTGTGATAATATA 846

Db 210 AlaThrArgValGlnAspAlaPheAlaAlaAlaLysLeuAlaLeuAlaLeuIle 229

QY 847 GTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCGCAGCACTTTAAAGACGCCCTTTTCA 906

Db 230 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 249

QY 907 -----GGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTAT 960

Db 250 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 269

QY 961 GCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAA 1020

Db 270 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 289

QY 1021 ACCATTCCCTTGCATATATGATATCCATGGCCATTGTGCACCATTTGGCTATGTGCTGACA 1080

Db 290 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr 309

QY 1081 AATGTGGCCTACTTTACGACCATTAATGCTGAGGAGTGTCTGCTTTTCAAAATGAGTGGCA 1140

Db 310 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 329

QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1200

Db 330 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 349

QY 1201 CTCTCTGCTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTCTATGTT 1260

Db 350 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 369

QY 1261 GCGTCTCGAGAGGGTCACTTCCAGAAATCCCTCCATGATTCATGTCGCGAAGCACACT 1320

Db 370 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 389

QY 1321 CCTCTACCAGCTGTTATTGTTTGGACCCCTTGGACAATGATAATGCTCTCTCTGGAGAC 1380

Db 390 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 409

QY 1381 CTCGACAGCTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGCTGGCAGTT 1440

Db 410 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 429

QY 1441 GCTGGCGTGATTTATCTTCGATACAAATGCCAGATATGTCATCGTCTCTTCAAGGTGCCA 1500

Db 430 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 449

QY 1501 CTGTTTCATCCAGCTTTGTTTTCCTTTCATGCCTTTCATGGTTGCCCTTTCCTCTAT 1560

Db 450 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 469

QY 1561 TCGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTAT 1620

Db 470 LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr 489

QY 1621 TATCTCTTTATTATATGGGACAAGAAACCCAGGTGGTTTAGAATAATGTGAGAGAAAATA 1680

Db 490 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 509

QY 1681 ACCAGAACATTACAAATAATACTGGAAGTTGTACCAAGAA 1722

Db 510 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 523

RESULT 5

US-10-163-866-38

; Sequence 38, Application US/10163866

; Publication No. US20030027188A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/10/163,866

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: US 60/357,600

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 38

; LENGTH: 507

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-163-866-38

Alignment Scores:

Pred. No.:	2,32e-108	Length:	507
Score:	1198.50	Matches:	231
Percent Similarity:	65.79%	Conservative:	94
Best Local Similarity:	46.76%	Mismatches:	166
Query Match:	29.09%	Indels:	3
DB:	9	Gaps:	2

US-09-667-170A-440 (1-2239) x US-10-163-866-38 (1-507)

QY 247 CCTTCCCTGGGCAACAAGGAGCCACCTGGGAGGAGCCCTTTTCAGGAAGAGAGCGCCTTT 306

Db 14 ProValAlaGluGluLysGluAlaArgGluLysMetLeuAlaSerLysArgAlaAsp 33

QY 307 TCAGGAAGAGAGCGCCTTTTCAGGAAGAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTA 366

Db 34 GlyAlaAlaProAlaGlyGluGly---GluGlyValThrLeuGlnArgAsnIleThrLeu 52

QY 367 CTGAGGGGAGTCTCCATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCATCTCTCCT 426

Db 53 LeuAsnGlyValAlaIleIleValGlyAlaIleIleGlySerGlyIlePheValThrPro 72

QY 427 AAGGGCGTGCTCCAGAACACGCGGCGGCGTGGGCATGCTCTGTGACCATCTGGACGGTGTGT 486

Db 73 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValMetTrpAlaAlaCys 92
QY 487 GGGGTCCTGTCACTATTGGAGCTTTGTCTTATGTCTGAATTGGGACAACTATAAGAAA 546
Db 93 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrIleSerLys 112
QY 547 TCTGGAGGTCAATACACATATATTTTGGAGTCTTTGGTCCATTTACCAGCTTTTGTACGA 606
Db 113 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCTGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTT 666
Db 133 LeuTrpIleGluLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATTTCTGGAACCATTTTTCATATGTGAATCCCTGAACTTGGCATCAAG 726
Db 153 AlaAlaTyrLeuLysProLeuPheProThrCysProValProGluAlaAlaLys 172
QY 727 CTCAATTACAGCTGTGGGCATAACTGTAGTGTATGCTGATGCTTAAATAGCATGTGTGAGTGT 786
Db 173 LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla 192
QY 787 AGCGCCGGATCCAGATTTTCTTAACCTTTTTCGAGCTCACAGCAATTCGTATAATTATA 846
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAlaIle 212
QY 847 GTCCCTGGAGTTATGCAGCTAATTAAGGTCAACCGCAGAACTTTAAAGACGCTTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGAAGAGATTCAAGTATTACGCGGTTGCCACCTGGCTTTTATTATGGAATGTAT 960
Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
QY 961 GCATATGCTGGTGGTGTTCCTCAACTTTTACTGAAGAGTAGAAAACCCCTGAAAAA 1020
Db 253 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACCATTCCCTTGCATATATGTATATCCATGGCCATTGTCACCATTTGGCTATGTGTGACA 1080
Db 273 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
QY 1081 AATGTGGCTTACTTTTACGACCATTAATGCTGAGGAGCTGTGCTTCAATGCAGTGGCA 1140
Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTCCGATCTTTTGTGCC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
QY 1201 CTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTCTATGTT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 352
QY 1261 GCGTCTCGAGAGGCTCACCTTCCAGAAATCTCTCCATGATTTCATGTCCGCAAGCACACT 1320
Db 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
QY 1321 CCTCTACAGCTGTATTGTTTTCACCCCTTTGACCAATGATAATGCTCTTCTCTGGAGAC 1380
Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
QY 1381 CTCGACAGCTTTTGAATTTCTCAGTTTTCAGGTTGGCTTTTATTGGCTGGCAGTT 1440
Db 393 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 412
QY 1441 GCTGGGCTGATTTATCTCGATACAAAATGCCAGATATGCATCGTCTTCAAGGTGCCA 1500
Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
QY 1501 CTGTTTCATCCAGCTTTGTTTCTTCCATGCTCTTTCATGGTGGCTTCCCTCTAT 1560
Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452

QY 1561 TCGGACCCATTTAGTACAGGATTGGCTTCGTCACTACTCTGACTGGAGTCCCTGCGTAT 1620
Db 453 LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr 472
QY 1621 TATCTCTTTTATATATGAGCAAGAAACCCAGGTGGTTTAGAATAATGTCAGAGAAATA 1680
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 492
QY 1681 ACCAGAACATTACAAATAATACTGGAAGTTGTACCAGAAAGAA 1722
Db 493 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 506
RESULT 6
US-10-163-866-39
; Sequence 39, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-39
Alignment Scores:
Pred. No.: 2,32e-108 Length: 507
Score: 1198.50 Matches: 231
Percent Similarity: 65.79% Conservative: 94
Best Local Similarity: 46.76% Mismatches: 166
Query Match: 29.09% Indels: 3
DB: 9 Gaps: 2
US-09-667-170A-440 (1-2239) x US-10-163-866-39 (1-507)

QY 247 CCTTCCCTGGGCAACAAGAGCCACCTGGGAGGACGCCTTTTCAGGAAGAGAGCGCCTTT 306
Db 14 ProValAlaGluGluLysGluGluAlaArgGluLysMetLeuAlaSerLysArgAlaAsp 33
QY 307 TCAGGAAGAGAGCGCCTTTTCAGGAAGAGAGAGAGAGTGCAGCTGAAGAGGAAAGTCACCTTTA 366
Db 34 GlyAlaAlaProAlaGlyGluGly---GluGlyValThrLeuGlnArgAsnIleThrLeu 52
QY 367 CTGAGGGGAGTCTCCATTCATTCATTTGGCACCACCATCATTCAGGAGCAGGAATCTTCATCTCT 426
Db 53 LeuAsnGlyValAlaIleIleValGlyAlaIleIleGlySerGlyIlePheValThrPro 72
QY 427 AAGGGCGTGTCTCCAGAACACACGCGGCGTGGGCGTGGGCGTGTCTGTACCATCTGGACGGTGTGT 486
Db 73 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValMetTrpAlaAlaCys 92
QY 487 GGGTCTCTGTCACCTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACTATATAAGAAA 546
Db 93 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleSerLys 112
QY 547 TCTGAGGTCATTACACATATATTTTGGAGCTCTTTGGTCCATTACCAGCTTTTGTACGA 606
Db 547 TCTGAGGTCATTACACATATATTTTGGAGCTCTTTGGTCCATTACCAGCTTTTGTACGA 606

Db 113 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCTGGGTGGAACCTCCTCATATAACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTT 666
Db 133 LeuTrpIleGluLeuLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATTCTGGAAACCAATTTTATTCAATGTGAATCCCTGAACTTGGCATCAAG 726
Db 153 AlaAlaTyrLeuLeuLysProLeuPheProThrCysProValProGluAlaAlaLys 172
QY 727 CTCATTACAGCTGTGGGCATAACTGTAGTGATGGTCCCTAAATAGCATGAGTGTCAAGTGG 786
Db 173 LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla 192
QY 787 AGCGCCCGGATCCAGATTTTCTTAACCTTTTGCAAGCTCACAGCACTTTAAAGACGCCCTTTCA 906
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAlaLeuIleIle 212
QY 847 GTCCCTGGAGTTATGCAGCTAATTAAGGTCAAAACGCCAGCACTTTAAAGACGCCCTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGAAGAGATTCAAGTATTACCGGGTTGCCACTGGCTTTTATTATGGAATGTAT 960
Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
QY 961 GCATATGCTGGTGGTGTTCCTCACTAATTTGTTACTGAAGAGTAGAAAAACCTTGAAAAA 1020
Db 253 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACCATTCCCTGCAATATGTATATCCCATGGCCATTGTCACCATTTGGCTATGTCTGACA 1080
Db 273 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
QY 1081 AATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAATGTCAGTGGCA 1140
Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
QY 1201 CTCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTATTCTATGTT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 352
QY 1261 GCGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCAATGTCGCGCAAGCACACT 1320
Db 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
QY 1321 CCTTACCAGCTGTTATTGTTTGGACCCCTTGCACCAATGATAATGCTCTTCTCTGGAGAC 1380
Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
QY 1381 CTCGACAGCTTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTTTATTGGGCTGGCAGTT 1440
Db 393 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 412
QY 1441 GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCCCTTTCAAGGTGCCA 1500
Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
QY 1501 CTGTTTCATCCAGCTTTGTTTTCCTTCCATGCTCTTCCATGCTGGCTTTCCCTCTAT 1560
Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452
QY 1561 TCGGACCCATTAGTACAGGATTGGCTTCGTCATCATCTGACTGGAGTCCCTCGGTAT 1620
Db 453 LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr 472
QY 1621 TATCTCTTTATATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTCCAGAGAAAAATA 1680
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 492

QY 1681 ACCAGAACATTACAAAATAATACTGGAAAGTTGTACCAGAAGAA 1722
Db 493 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 506
RESULT 7
US-10-163-866-40
; Sequence 40, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-40
Alignment Scores:
Pred. No.: 3e-104 Length: 515
Score: 1156.50 Matches: 218
Percent Similarity: 69.16% Conservatives: 105
Best Local Similarity: 46.68% Mismatches: 141
Query Match: 28.07% Indels: 3
DB: 9 Gaps: 2
US-09-667-170A-440 (1-2239) x US-10-163-866-40 (1-515)
QY 334 GAGAAAGTGCAGCTGAAGAGGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 37 GluThrMetGlnLeuLysLysGluIleSerLeuLeuAsnGlyValSerLeuValGly 56
QY 394 ACCATCATTTGGACGAGGAATCTTCTCATCTCTCTTAAGGGCTGCTCCAGAACACGGCAGC 453
Db 57 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuValHisThrAlaSer 76
QY 454 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGTCTGTGCTACTATTTGGAGCTTTG 513
Db 77 TyrGlyMetSerLeuIleValTrpAlaIleGlyGlyLeuPheSerValValGlyAlaLeu 96
QY 514 TCTTATGCTGAATTTGGGAACAACACTATAAAGAAATCTGGAGGTCAATACACATATATTTG 573
Db 97 CysTyrAlaGluLeuGlyThrThrIleThrLysSerGlyAlaSerTyrAlaTyrIleLeu 116
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAATACGC 633
Db 117 GluAlaPheGlyGlyPheIleAlaPheIleArgLeuTrpValSerLeuLeuValValGlu 136
QY 634 CCTGCACCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATTTCTGGACCACTTTT 693
Db 137 ProThrGlyGlnAlaIleIleAlaIleThrPheAlaAsnTyrIleIleGlnProSerPhe 156
QY 694 ATTCATGTGAATCCCTGAACCTTCCGATCAAGCTCAATACAGCTGTGGCATAACTGTA 753
Db 157 ProSerCysAspProProTyrLeuAlaCysArgLeuLeuAlaAlaCysIleCysLeu 176
QY 754 GTGATGGTCTCTAAATAGCATGAGTGTGAGCTGGAGCGCCCGGATTTCTTAACC 813

FILE REFERENCE: EX02-080C
CURRENT APPLICATION NUMBER: US/10/163,866
CURRENT FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: US 60/296,076
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/328,605
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/338,733
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: US 60/357,253
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/357,600
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 511
TYPE: PRT
ORGANISM: Homo sapiens
US-10-163-866-41

Alignment Scores: 8.36e-102 Length: 511
Pred. No.: 1131.50 Matches: 214
Score: 68.00% Conservative: 109
Percent Similarity: 45.05% Mismatches: 149
Best Local Similarity: 27.46% Indels: 3
Query Match: 9 Gaps: 2
DB:

US-09-667-170A-440 (1-2239) x US-10-163-866-41 (1-511)

QY 310 GGAAGAGACGCGCTTTTCAGGAAGAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTG 369
Db 21 GlyAspGlyAlaSerProGlyProGluGlnValLysLeuLysLysGluLeuSerLeuLeu 40
QY 370 AGGGAGTCTCCATTATCATTTGGACCATCATTTGGAGCAGGAATCTTCATCTCTCTTAAG 429
Db 41 AsnGlyValCysLeuIleValGlyAsnMetIleGlySerGlyIlePheValSerProLys 60
QY 430 GCGTGTCTCCAGAACACGGGCGGCGTGGCATGTCTCTGACCATCTGGACGGTGTGTGG 489
Db 61 GlyValLeuIleTyrSerAlaSerPheGlyLeuSerLeuValIleTrpAlaValGlyGly 80
QY 490 GTCCTGTCTACTATTTGGAGCTTTGTCTTATGTCTGAATTTGGGACAACTATAAGAAATCT 549
Db 81 LeuPheSerValPheGlyAlaLeuLeuGlyAlaGluLeuGlyThrThrIleLysLysSer 100
QY 550 GGAGGTCAATACACATATATTTTGGAGCTTTTGGTCCATTACAGCTTTTGTACGAGTC 609
Db 101 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 120
QY 610 TGGTGGAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTGTATCCCTGGCATTGGA 669
Db 121 TrpThrSerLeuLeuIleIleGluProThrSerGlnAlaIleIleAlaIleThrPheAla 140
QY 670 CGCTACATCTGGAACCATTTTATTCAATGTGAAATCCCTGAACCTGCGATCAAGCTC 729
Db 141 AsnTyrMetValGlnProLeuPheProSerCysPheAlaProTyrAlaAlaSerArgLeu 160
QY 730 ATTACAGCTGTGGGCATACCTAGTGTAGTGTGTCCTAAATAGCATGAGTGTGAGTGGAGC 789
Db 161 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly 180
QY 790 GCCCGGATCCAGATTTTCTTAACCTTTTTCAGAGCTCACAGCAATTTCTGATAATTATAGTC 849
Db 181 ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIleVal 200
QY 850 CCTGGAGTTATGCAGCTAATTAAGGTCAAACGCGAGAACTTTAAAGACGCCCTTTTCAGGA 909
Db 201 AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly 220
QY 910 AGAGATTCAAGTATTACGCGGTGGCCACTGGCTTTTATTATGGAATGTAATGATATGCT 969

Db 177 LeuThrPheValAsnCysAlaTyrValLysTrpGlyThrArgValGlnAspThrPheThr 196
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATAA 873
Db 197 TyrAlaLysValValAlaLeuIleAlaIleIleValMetGlyLeuValLysLeuCysGln 216
QY 874 GGTCAACGCGAGAACTTTAAAGACGCGCTTTTCAGGAAGAGATTCAGTATTACGCGGTTG 933
Db 217 GlyHisSerGluHisPheGlnAspAlaPheGluGlySerSerTyrPheMetGlyAsnLeu 236
QY 934 CCACTGGCTTTTATATGGAATGTATGCATATGCTGGCTGGTGTTCCTCAACCTTTGTT 993
Db 237 SerLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTyrPheThrLeuAsnPheVal 256
QY 994 ACTGAAGAAGTAGAAACCCCTGAAAAACCATTCCTTGCATATATGATATCCATGGCC 1053
Db 257 ThrGluGluIleLysAsnProGluArgAsnLeuProLeuAlaIleGlyIleSerMetPro 276
QY 1054 ATTGTCAACCATTTGGCTATGTGTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 1113
Db 277 IleValThrLeuIleTyrIleLeuThrAsnValAlaTyrTyrThrValLeuAsnIleSer 296
QY 1114 GAGTGTCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 297 AspValLeuSerSerAspAlaValAlaValThrPheAlaAspGlnThrPheGlyMetPhe 316
QY 1174 TCATTAGCAGTTCGATCTTGTGTCCTCTCTCCCTCTCTGCTGCTTGGCTCCATGAACGGTGTG 1233
Db 317 SerTrpThrIleProIleAlaValAlaLeuSerCysPheGlyGlyLeuAsnAlaSerIle 336
QY 1234 TTTGCTGTCTCCAGGTTATTTCTATGTTGCTCTCGAGAGGGTCACTTCCAGAAATCCTC 1293
Db 337 PheAlaSerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProAspLeuLeu 356
QY 1294 TCCATGATTCATGTCGCAAGCACACTCTCTACACAGTCTTTTGAATTTTCCCTCAGTTTGC 1413
Db 357 SerMetIleHisIleGluArgPheThrProIleProAlaLeuLeuPheAsnCysThrMet 376
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTTCCCTCAGTTTGC 1413
Db 377 AlaLeuIleTyrLeuIleValGluAspValPheGlnLeuIleAsnTyrPheSerPheSer 396
QY 1414 AGGTGGCTTTTATTTGGCTGGCAGTCTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 397 TyrTrpPhePheValGlyLeuSerValValGlyGlnLeuTyrLeuArgTrpLysGluPro 416
QY 1474 GATATGCATGCTCTTCAAGGTGCACACTGTTTCATCCCGAGCTTTGTTTTCCTTCACATGC 1533
Db 417 LysArgProArgProLeuLysLeuSerValPhePheProIleValPheCysIleCysSer 436
QY 1534 CTCTTCATGTTGCTCTTCCCTCTATTCGGACCCATTTACTACAGGATTTGGCTTCGTC 1593
Db 437 ValPheLeuValIleValProLeuPheThrAspThrIleAsnSerLeuIleGlyIleGly 456
QY 1594 ATCACTCTGAGTGGAGTCCCTGCGTATATCTC-----TTTATTATATGGGACAGAAA 1647
Db 457 IleAlaLeuSerGlyValProPheTyrPheMetGlyValTyrLeuProGluSerArgArg 476
QY 1648 CCAAGGTGGTTTAGAATAATGTACAGAGAAATAAACAGAGAGAGAGAGAGAGAGAGAG 1704
Db 477 ProLeuPheIleArgAsnValLeuAlaAlaIleThrArgGlyThrGlnGlnLeuCysPhe 496
QY 1705 GAAGTGTACCAAGAGAGAT 1725
Db 497 CysValLeuThrGluLeuAsp 503

RESULT 8
US-10-163-866-41
; Sequence 41, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

Db	281	TyTyrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe	300
QY	1150	TCTGAGCGGCTACTGGGAAATTTCTATTAGCAGTTCCGATCTTTGTTGCCCTCTCTCCTGC	1209
Db	301	AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys	320
QY	1210	TTTGGCTCCATGAACGGTGGTGTGTTTGGTCTCTCCAGGTTATTCTATGTTGCGTCTCGA	1269
Db	321	PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuPhePheValGlySerArg	340
QY	1270	GAGGGTCACCTTCAGAAATCCTCTCCATGATTCATGTCGCGAAGCACACTCTCTACCA	1329
Db	341	GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro	360
QY	1330	GCTGTTATTGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTCGACAGT	1389
Db	361	SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln	380
QY	1390	CTTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTG	1449
Db	381	LeuIleAsnTyTyrSerPheSerTyTyrPhePheValGlyLeuSerIleValGlyGln	400
QY	1450	ATTTATCTTCGATACAAATGCCAGATATGCATCGTCCTTTTCAAGGTGCCACTGTTTCATC	1509
Db	401	LeuTyTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe	420
QY	1510	CCAGCTTTGTTTTCCTTCACATGCCTCTTTCATGGTTGCCCTTTCCCTCTATTCCGACCCA	1569
Db	421	ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyTyrSerAspThr	440
QY	1570	TTTAGTACAGGGATTGGCTTCGTTCATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTT	1629
Db	441	IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyTyrPheLeuIle	460
QY	1630	ATT-----ATATGGGACAGAAACCCAGGTGGTTTAGAATAATGTCAGAGAAATAACC	1683
Db	461	IleArgValProGluHisLysArgProLeuTyTyrLeuArgGlnIleValGlySerAlaThr	480
QY	1684	AGAACATTACAAATAATA---CTGGAAGTTGTACCAGAGAAGAT	1725
Db	481	ArgTyTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp	495

```

RESULT 10
US-10-163-866-43
; Sequence 43, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7S AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 535
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-163-866-43

```

Alignment Scores:		
pred. No.:	7.35e-99	Length:
Score:	1101.50	Matches:
		535
		220

Percent Similarity:	65.52%	Conservative:	103
Best Local Similarity:	44.62%	Mismatches:	161
Query Match:	26.74%	Indels:	9
DB:	9	Gaps:	4
US-09-667-170A-440 (1-2239) x US-10-163-866-43 (1-535)			
QY	262	AAGGAGCCACGTGGCAGGACGCCCTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGACGCC	321
DB	13	LysLysHisProGly-----GlyGlyGluSerAspAlaSerProGluAla	27
QY	322	TTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTACCTTACTGAGGGAGTCTCC	381
DB	28	GlySerGlyGlyGlyValAlaLeuLysLysGluIleGlyLeuValSerAlaCysGly	47
QY	382	ATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCCAG	441
DB	48	IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu	67
QY	442	AACACGGGACGGTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTCACTA	501
DB	68	AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyPheIleThrVal	87
QY	502	TTTGGAGCTTTTGCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTAC	561
DB	88	ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr	107
QY	562	ACATATATTTGGAACTCTTTGGTCCATTTACCAGCTTTTGTACGAGTCTGGTGGAACTC	621
DB	108	SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal	127
QY	622	CTCATAAATACGCCCTCCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTCTG	681
DB	128	LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu	147
QY	682	GAACCATTTTATTCAATGTGAAATCCCTGAACTTGGCATCAAGCTCATTCACAGCTGTG	741
DB	148	GlnProLeuPheProThrCysPheProProGluSerGlyLeuArgLeuLeuAlaIle	167
QY	742	GGCATAACTGTAGTGATGGTCCTAAATAGCATGAGTGTGAGCTGGAGCGCCCGGATCCAG	801
DB	168	CysLeuLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln	187
QY	802	ATTTTCTTAACCTTTTGCAGGCTCACAGCAATCTGATAATTATAGTCCCTGGAGTTATG	861
DB	188	AspIlePheThrAlaGlyLysLeuLeuAlaLeuAlaLeuIleIleMetGlyIleVal	207
QY	862	CAGCTAATTAAGGTCAA-----ACGCAGAACTTTAAAGACGCCCTTTTCAGGA---AGA	912
DB	208	GlnIleCysLysGlyGluTyrPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln	227
QY	913	GATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATGCATATGCTGGC	972
DB	228	GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaTyrGlyGly	247
QY	973	TGGTTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAACCAATTCCTCCCTT	1032
DB	248	TrpAsnPheLeuAsnTyrValThrGluLeuValAspProTyrLysAsnLeuProArg	267
QY	1033	GCAATATGATATCCATGGCCATTGTCCACCATTTGGCTATGTGCTGACAAATGTGGCCTAC	1092
DB	268	AlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsnValAlaTyr	287
QY	1093	TTTACGACCATTAATGCTGAGGAGCTGCTGCTTTTCAATTCAGTGGCAGTGACCTTTCT	1152
DB	288	ValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaValThrPheGly	307
QY	1153	GAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTT	1212
DB	308	GluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeuSerThrPhe	327
QY	1213	GGCTCCCATGAACGGTGGTGTGTTTGTGCTCTCCAGGTTATTCTATGTTGCGTCTCGAGAG	1272

Db 328 GlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGlyAlaArgGlu 347
QY 1273 GGTCACTTCCAGAAATCCTCTCCATGATTTCATGTCGCAAGCACACTCTCTACCAGCT 1332
Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrProIleProAla 367
QY 1333 GTTATTGTTTGCACCCCTTTGACAATGATAATGCTCTCTCTCTGGAGACCTCGACAGTCTT 1392
Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu 387
QY 1393 TTGAATTCCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTGTGGGTGATT 1452
Db 388 IleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle 407
QY 1453 TATCTCGATACAAATGCCAGATATGCATCGCTCTTCAAGGTGCCACTGTTTCATCCCA 1512
Db 408 ValLeuArgTrpLysLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
QY 1513 GCTTGTGTTTCTTCACATGCCTCTTTCATGGTGGCCCTTTCCCTCTCTATTCGGACCCATT 1572
Db 428 IleIleTyrLeuLeuPheTrpAlaPheLeuLeuValPheSerLeuTrpSerGluProVal 447
QY 1573 AGTACAGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATT 1632
Db 448 ValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPheLeuGlyVal 467
QY 1633 ATATGGGACAAGAAACCCAGGTGGTGTAGATAATGTACAGAGAAATAACCAAGACATTA 1692
Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuThrLeuValSer 487
QY 1693 CAAATAATACTGGAAGTTGTA---CCAGAAGAAGATAAG 1728
Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500

RESULT 11

US-10-163-866-44
; Sequence 44, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 535
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-163-866-44

Alignment Scores:
Pred. No.: 7.35e-99 Length: 535
Score: 1101.50 Matches: 220
Percent Similarity: 65.52% Conservative: 103
Best Local Similarity: 44.62% Mismatches: 161
Query Match: 26.74% Indels: 9
DB: 9 Gaps: 4

US-09-667-170A-440 (1-2239) x US-10-163-866-44 (1-535)

QY 262 AAGGAGCCACCTGGGACGACGCTTTTCAGGAAGAGACGCTTTTCAGGAAGAGACGCC 321

Db 13 LysLysHisProGly-----GlyGlyGluSerAspAlaSerProGluAla 27
QY 322 TTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCC 381
Db 28 GlySerGlyGlyGlyValAlaLeuLysLysGluIleGlyLeuValSerAlaCysGly 47
QY 382 ATTATCATTTGGCACCACATTTGGAGCAGGAATCTTTCATCTCTCTCTAAGGGCTGCTCCAG 441
Db 48 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 67
QY 442 AACACGGGCGAGCTGGGCATGTCTCTGACCATCTGGACGGTGTGGGTCTCTGTCACCTA 501
Db 68 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyPheIleThrVal 87
QY 502 TTTGGAGCTTTGCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCACTAC 561
Db 88 ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 107
QY 562 ACATATATTTTGAAGTCTTTTGGTCATTTACCAGCTTTTGTACGAGTCTGGGTGGAATC 621
Db 108 SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 127
QY 622 CTCATAATACGCCCTGCAGCTACTCTGTGATATCCCTGGCATTTGGACGCTACATCTG 681
Db 128 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 147
QY 682 GAACCATTTTTCATCAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTCACAGCTGTG 741
Db 148 GlnProLeuPheProThrCysPheProProGluSerGlyLeuArgLeuAlaIle 167
QY 742 GGCATAACTGTAGTGTGTCCTCTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAG 801
Db 168 CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 187
QY 802 ATTTTCTTAACCTTTTTCAGAGCTCACAGCAATTTCTGATAAATATAGTCCCTGGAGTTATG 861
Db 188 AspIlePheThrAlaGlyLysLeuLeuAlaLeuIleIleIleMetGlyIleVal 207
QY 862 CAGCTAATTAAGGTCAA-----ACGCAGAACTTTAAAGACGCCCTTTTCAGGA---AGA 912
Db 208 GlnIleCysLysGlyGluTyrPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln 227
QY 913 GATTCAAGTATTACGGGTTGCCACTGGCTTTTATTTATTTATGGAATGTATGCATATGCTGC 972
Db 228 GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaTyrGlyGly 247
QY 973 TGTGTTTACCTCAACTTTGCTTACTGAAGAAGTAGAAACCCCTGAAACCACTTCCCTT 1032
Db 248 TrpAsnPheLeuAsnTyrValThrGluGluLeuValAspProTyrLysAsnLeuProArg 267
QY 1033 GCAATATGTATATCCATGGCCATTTGTCCACCATTTGGCTATGTGCTGACAAATGTGGCCTAC 1092
Db 268 AlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsnValAlaTyr 287
QY 1093 TTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCT 1152
Db 288 ValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaValThrPheGly 307
QY 1153 GAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTTGTTGCCCTCTCTGCTTT 1212
Db 308 GluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeuSerThrPhe 327
QY 1213 GGCTCCATGAACGGTGTGTTTGTGCTGCTCCAGGTTATTCATGTTGCGTCTCGAGAG 1272
Db 328 GlyGlyValAsnGlySerLeuPheThrSerArgLeuPhePheAlaGlyAlaArgGlu 347
QY 1273 GGTCACTTCCAGAAATCCTCTCCATGATTCATGTCCCGCAAGCACACTCTCTACCAGCT 1332
Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrProIleProAla 367
QY 1333 GTTATTGTTTTCACCCCTTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTT 1392


```
Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu 387
QY 1393 TTGAATTCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATT 1452
Db 388 ileAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle 407
QY 1453 TATCTTCGATACAAATGCCAGATATGTCATCGTCTTCAAGGTGCCACTGTTTCATCCCA 1512
Db 408 ValLeuArgTrpLysLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
QY 1513 GCTTTGTTTCTTCACATGCCCTTCATGGTTGGCTTCCCTTCCCTCTATTTCGGACCCATT 1572
Db 428 IleIleTyrLeuLeuPheTrpAlaPheLeuLeuValPheSerLeuTrpSerGluProVal 447
QY 1573 AGTACAGGATGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTATT 1632
Db 448 valCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPheLeuGlyVal 467
QY 1633 ATATGGGACAGAAACCCAGGTGGTTTGAATAATGTCCAGAGAAATAACAGAACATTA 1692
Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuLeuThrLeuValSer 487
QY 1693 CAAATAATCTGGAAGTTGTA---CCAGAAGAAGATAAG 1728
Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500

RESULT 12
US-10-163-866-46
; Sequence 46, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-46

Alignment Scores:
Pred. No.: 2,47e-96 Length: 487
Score: 1075.50 Matches: 207
Percent Similarity: 66.16% Conservative: 98
Best Local Similarity: 44.90% Mismatches: 153
Query Match: 26.10% Indels: 3
DB: 9 Gaps: 2

US-09-667-170A-440 (1-2239) x US-10-163-866-46 (1-487)
QY 346 CTGAACAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTTGGCCACCATTGGA 405
Db 26 LeuGlnLysGluLeuGlyLeuIleSerGlyIleSerIleIleValGlyThrIleGly 45
QY 406 GCAGGAATCTTCATCTCTCCTAAGGGCGTGTCCAGAACACGGCAGCGTGGGCATGTCT 465
Db 46 SerGlyIlePheValSerProLysSerValLeuSerAsnThrGluAlaValGlyProCys 65
QY 466 CTGACCATCTGGACGGGTGTGGGGTCTGTCACTATTGGAGCTTTGTCTTATGCTGAA 525
```

```
Db 66 LeuIleIleTrpAlaAlaCysGlyValLeuAlaThrLeuGlyAlaLeuCysPheAlaGlu 85
QY 526 TTGGGAACAACATAAAAGAAATCTGAGGTCATTACACATATATTTTGGAAAGTCTTTGGT 585
Db 86 LeuGlyThrMetIleThrLysSerGlyGlyGluTyrProTyrLeuMetGluAlaTyrGly 105
QY 586 CCATTACCAGCTTTGTACGAGTCTGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACT 645
Db 106 ProIleProAlaTyrLeuPheSerTrpAlaSerLeuIleValIleLysProThrSerPhe 125
QY 646 GCTGTGATATCCCTGGCATTGTGGACGCTACATTCCTGGAACCATTTTATTCAATGTGAA 705
Db 126 AlaIleIleCysLeuSerPheSerGluTyrValCysAlaProPheTyrValGlyCysLys 145
QY 706 ATCCCTGAACCTTCGATCAAGCTCAATTACAGCTGTGGCATACCTGTAGTGTGTCCTA 765
Db 146 ProProGlnIleValValLysCysLeuAlaAlaAlaIleLeuPheIleSerThrVal 165
QY 766 AATAGCATGAGTGTGAGCTGGAGCGCCGCGATCCAGATTTTCTTAACCTTTTGAAGCTC 825
Db 166 AsnSerLeuSerValArgLeuGlySerTyrValGlnAsnIlePheThrAlaAlaLysLeu 185
QY 826 ACAGCAATCTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCGAG 885
Db 186 ValIleValAlaIleIleIleSerGlyLeuValLeuLeuAlaGlnGlyAsnThrLys 205
QY 886 AACTTTAAAGACGCCCTTTTCAGGAAGAGATCAAGTATTACGGGTTCGCCACTGGCTTTT 945
Db 206 AsnPheAspAsnSerPheGluGlyAlaGlnLeuSerValGlyAlaIleSerLeuAlaPhe 225
QY 946 TATTATGGAATGTATGCATATGCTGGCTGGTGTTCCTCAACTTGTGTACTGAAGAAGTA 1005
Db 226 TyrAsnGlyLeuTrpAlaTyrAspGlyTrpAsnGlnLeuAsnTyrIleThrGluGluLeu 245
QY 1006 GAAAACCCCTGAAAACCATTCCTCCCTTGCATATGTATATCCATCCATGGCCATTGTCA 1065
Db 246 ArgAsnProTyrArgAsnLeuProLeuAlaIleIleIleGlyIleProLeuValThrAla 265
QY 1066 GGCTATGTGCTGACAAATGTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTT 1125
Db 266 CysTyrIleLeuMetAsnValSerTyrPheThrValMetThrAlaThrGluLeuGln 285
QY 1126 TCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTT 1185
Db 286 SerGlnAlaValAlaValThrPheGlyAspArgValLeuTyrProAlaSerTrpIleVal 305
QY 1186 CCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCC 1245
Db 306 ProLeuPheValAlaPheSerThrIleGlyAlaAlaAsnGlyThrCysPheThrAlaGly 325
QY 1246 AGGTTATTCTATGTTGCGTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCAT 1305
Db 326 ArgLeuIleTyrValAlaGlyArgGluGlyHisMetLeuLysValLeuSerTyrIleSer 345
QY 1306 GTCCGCAAGCACACTCCTCTACAGCTGTTATTGTTTTCACCCCTTTGACAAATGATAATG 1365
Db 346 ValArgArgLeuThrProAlaProAlaIleIlePheTyrGlyIleAlaThrIleTyr 365
QY 1366 CTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTT 1425
Db 366 IleIleProGlyAspIleAsnSerLeuValAsnTyrPheSerPheAlaAlaTrpLeuPhe 385
QY 1426 ATTTGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATACAAATGCCAGATATGATCGT 1485
Db 386 TyrGlyLeuThrIleLeuGlyLeuIleValMetArgPheThrArgLysGluLeuArg 405
QY 1486 CCTTTCAAGGTGCCACTGTTTCATCCAGCTTTTGTGTTTTCCTTCACATGCCTCTTCATGTT 1545
Db 406 ProIleLysValProValValIleProValLeuMetThrLeuIleSerValPheLeuVal 425
QY 1546 GCCCTTTCCCTCTATTTCGGACCCCTATTAGTACAGGATTTGGCTTCGTCATC---ACTCTG 1602
```


Db 426 LeuAlaProIleIleSerLysProThrTrpGluTyrLeuTyrCysValLeuPheIleLeu 445
 QY 1603 ACTGGAGTCCCTGCGTATTATCTCTTATTATATGGGACAAAGAACCCAGGTGGTTAGA 1662
 Db 446 SerGlyLeuLeuPheTyrPheLeuPheValHisTyr-----LysPheGlyTrpAlaGln 463
 QY 1663 ATAATGTCAGAGAAAATAACAGAACATTACAAATAATACTGGAAGTTGTACCAGAAGAA 1722
 Db 464 LysIleSerLysProIleThrMetHisLeuGlnMetLeuMetGluValValProGlu 483
 QY 1723 GAT 1725
 Db 484 Glu 484
 RESULT 13
 US-10-163-866-45
 ; Sequence 45, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: US 60/296,076
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 45
 ; LENGTH: 487
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-163-866-45
 Alignment Scores:
 Pred. No.: 1.5e-95 Length: 487
 Score: 1067.50 Matches: 206
 Percent Similarity: 65.94% Conservative: 98
 Best Local Similarity: 44.69% Mismatches: 154
 Query Match: 25.91% Indels: 3
 DB: 9 Gaps: 2
 US-09-667-170A-440 (1-2239) x US-10-163-866-45 (1-487)
 QY 346 CTGAAGAGGAAAGTCACTTACTAGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGA 405
 Db 26 LeuGlnLysGluLeuGlyLeuIleSerGlyIleSerIleValGlyThrIleIleGly 45
 QY 406 GCAGGAATCTCATCTCTCCTAAGGCGGTGCTCCAGAACACGGCGGTGGCATGTCT 465
 Db 46 SerGlyIlePheValSerSerLysSerValLeuSerAsnThrGluAlaValGlyProCys 65
 QY 466 CTGACCATCTGGACGGTGTCTGGGTCCTGTCTACTATTGGAGCTTTGTCTTATGCTGAA 525
 Db 66 LeuIleIleTrpAlaAlaCysGlyValLeuAlaThrLeuGlyAlaLeuCysPheAlaGlu 85
 QY 526 TTGGGAACAACATAAGAAATCTGGAGGTTCATTACACATATATTTTGAAGTCTTTGGT 585
 Db 86 LeuGlyThrMetIleThrLysSerGlyGlyGlyTyrProTyrLeuMetGluAlaTyrGly 105
 QY 586 CCATTACCAGCTTTTGTACGAGTCTGGTGGAACTCCTCATATAATACGCCCTGCAGCTACT 645
 Db 106 ProIleProAlaTyrLeuPheSerTrpAlaSerLeuIleValIleLysProThrSerPhe 125
 QY 646 GCTGTGATATCCCTGGCATTGGACGCTACATTCTGGGAACCATTTTTTATTCAATGTGAA 705

Db 126 AlaIleIleCysLeuSerPheSerGluTyrValCysAlaProPheTyrValGlyCysLys 145
 QY 706 ATCCCTGAAGTTCGATCAAGCTCATTACAGCTGTGGGCATTAACCTAGTAGTGATGTCCTA 765
 Db 146 ProProGlnIleValLysCysLeuAlaAlaAlaIleLeuPheIleSerThrVal 165
 QY 766 AATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTC. 825
 Db 166 AsnSerLeuSerValArgLeuGlySerTyrValGlnAsnIlePheThrAlaAlaLysLeu 185
 QY 826 ACAGCAATTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAAGGTCAAACGAG 885
 Db 186 ValIleValAlaIleIleIleSerGlyLeuValLeuAlaGlnGlyAsnThrLys 205
 QY 886 AACTTTAAAGACGCTTTTCAAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTT 945
 Db 206 AsnPheAspAsnSerPheGluGlyAlaGlnLeuSerValGlyAlaIleSerLeuAlaPhe 225
 QY 946 TATTATGGAATGTATGCATATGCTGGCTGGTGTTCCTCAACTTTTGTACTGAAGAAGTA 1005
 Db 226 TyrAsnGlyLeuTrpAlaTyrAspGlyTrpAsnGlnLeuAsnTyrIleThrGluGluLeu 245
 QY 1006 GAAACCCCTGAAAAACCAATTCCCTTGCAATATGATATATCCATGGCCATTGTACACATT 1065
 Db 246 ArgAsnProTyrArgAsnLeuProLeuAlaIleIleIleGlyIleProLeuValThrAla 265
 QY 1066 GGCTATGTCTGACAAATGTGGCTACTTTTACGACCATTAATGCTGAGGAGCTGTGCTT 1125
 Db 266 CysTyrIleLeuMetAsnValSerTyrPheThrValMetThrAlaThrGluLeuGln 285
 QY 1126 TCAATGTCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTT 1185
 Db 286 SerGlnAlaValAlaValThrPheGlyAspArgValLeuTyrProAlaSerTrpIleVal 305
 QY 1186 CCGATCTTTTGTGGCTCTCCTGCTTGGTCCCATGAACGGTGGTGTGTTGCTGTCTCC 1245
 Db 306 ProLeuPheValAlaPheSerThrIleGlyAlaAlaAsnGlyThrCysPheThrAlaGly 325
 QY 1246 AGTTATTCTATGTTGCTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTTCAT 1305
 Db 326 ArgLeuIleTyrValAlaGlyArgGluGlyHisMetLeuLysValLeuSerTyrIleSer 345
 QY 1306 GTCCGCAAGCACACTCTCTACAGCTGTATTGTTGTTTGCACCCCTTTGACAAATGATAATG 1365
 Db 346 ValArgArgLeuThrProAlaProAlaIleIlePheTyrGlyIleIleAlaThrIleTyr 365
 QY 1366 CTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTT 1425
 Db 366 IleIleProGlyAspIleAsnSerLeuValAsnTyrPheSerPheAlaTrpLeuPhe 385
 QY 1426 ATTGGCTGGCAGTTGCTGGCTGATTATTATCTTCGATACAAATGCCAGATATGATCGT 1485
 Db 386 TyrGlyLeuThrIleLeuGlyLeuIleValMetArgPheThrArgLysGluLeuGluArg 405
 QY 1486 CTTTCAAGGTGCCACTGTTCATCCCAGCTTTTGTTCCTTCACATGCCCTCTTTCATGTT 1545
 Db 406 ProIleLysValProValIleProValLeuMetThrLeuIleSerValPheLeuVal 425
 QY 1546 GCCCTTCCCTCTATTTCGGACCCCATTTTAGTACAGGGATTGGCTTCGTATC---ACTCTG 1602
 Db 426 LeuAlaProIleIleSerLysProThrTrpGluTyrLeuTyrCysValLeuPheIleLeu 445
 QY 1603 ACTGGAGTCCCTGCGTATTATCTCTTTTATTATATGGGACAAAGAACCCAGGTGGTTAGA 1662
 Db 446 SerGlyLeuLeuPheTyrPheLeuPheValHisTyr-----LysPheGlyTrpAlaGln 463
 QY 1663 ATAATGTCAGAGAAAATAACAGAACATTACAAATAATACTGGAAGTTGTACCAGAAGAA 1722
 Db 464 LysIleSerLysProIleThrMetHisLeuGlnMetLeuMetGluValValProGlu 483
 QY 1723 GAT 1725
 Db 484 Glu 484

;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: US 60/328,605
;; PRIOR FILING DATE: 2001-10-10
;; PRIOR APPLICATION NUMBER: US 60/338,733
;; PRIOR FILING DATE: 2001-10-22
;; PRIOR APPLICATION NUMBER: US 60/357,253
;; PRIOR FILING DATE: 2002-02-15
;; PRIOR APPLICATION NUMBER: US 60/357,600
;; PRIOR FILING DATE: 2002-02-15
;; NUMBER OF SEQ ID NOS: 54
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 47
;; LENGTH: 523
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-163-866-47

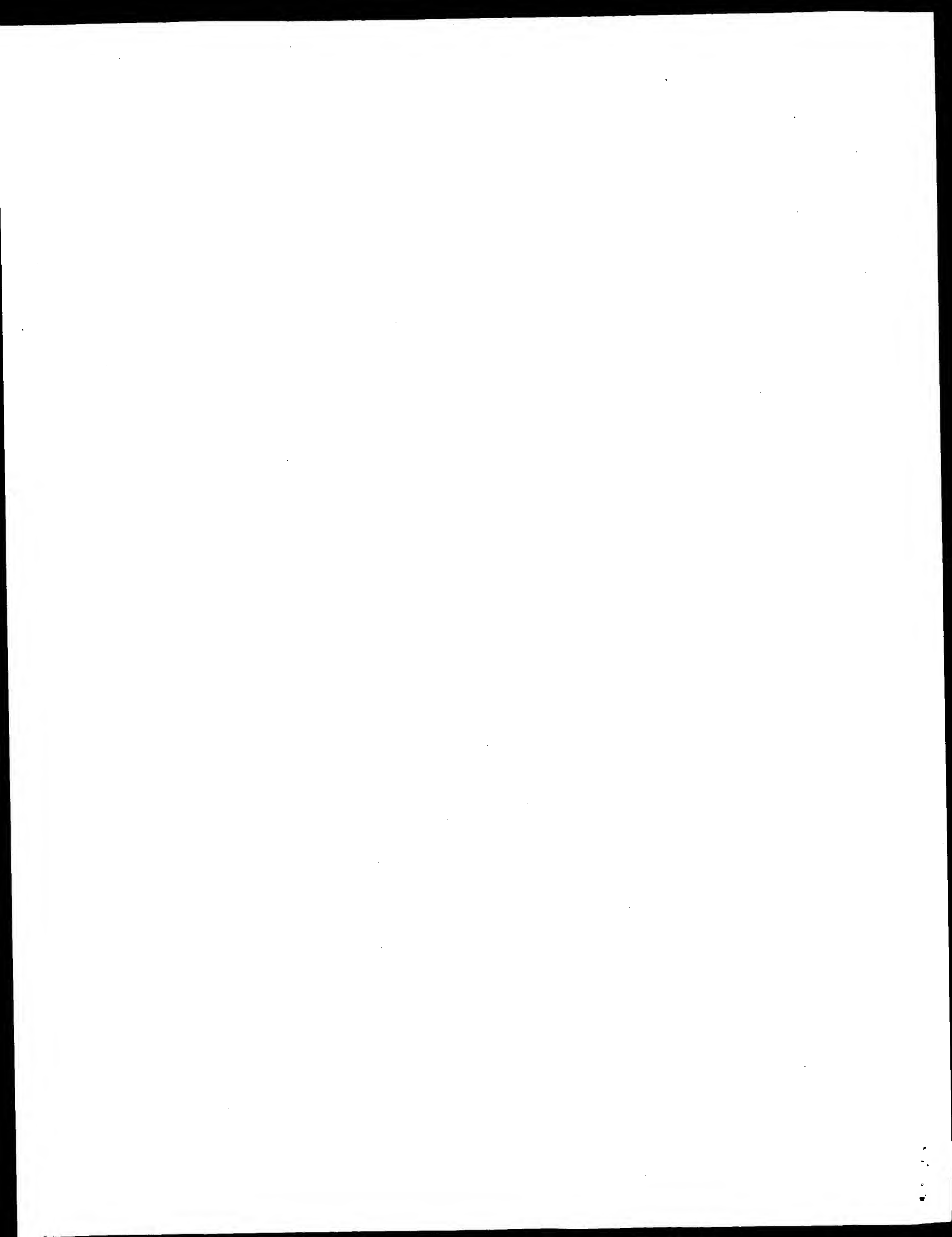
Alignment Scores:
Pred. No.: 9 48e-93 Length: 523
Score: 1039.00 Matches: 211
Percent Similarity: 64.40% Conservative: 102
Best Local Similarity: 43.42% Mismatches: 157
Query Match: 25.22% Indels: 16
DB: 9 Gaps: 5

US-09-667-170A-440 (1-2239) x US-10-163-866-47 (1-523)

QY 334 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
Db 32 GluArgValAlaLeuLysLysGluIleGlyLeuLeuSerAlaCysThrIleIleGly 51
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCCAGAACACGGGCAGC 453
Db 52 AsnIleIleGlySerGlyIlePheIleSerProLysGlyValLeuGluHisSerGlySer 71
QY 454 GTGGGCATGTCTGTGACCATCTGGACGGTGTGTGGGGTCTGTCACTATTTGGAGCTTTG 513
Db 72 ValGlyLeuAlaLeuPheValTrpValLeuGlyGlyValThraLeuGlySerLeu 91
QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCACTACACATATATTTG 573
Db 92 CysTyrAlaGluLeuGlyValAlaIleProLysSerGlyGlyAspTyrAlaTyrValThr 111
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCATCAATAACGC 633
Db 112 GluIlePheGlyGlyLeuAlaGlyPheLeuLeuLeuTrpSerAlaValLeuIleMetTyr 131
QY 634 CCTGCAGTACTGCTGTGATATCCCTGGCATTGAGAGGTACATCTGGAAACCATTTT 693
Db 132 ProThrSerLeuAlaValIleSerMetThrPheSerAsnTyrValLeuGlnProValPhe 151
QY 694 ATTCATGTGAATCCCTGAACCTTCCGATCAAGCTCATACAGCTGTGGGCATAACTGTA 753
Db 152 ProAsnCysIleProProThrThrAlaSerArgValLeuSerMetAlaCysLeuMetLeu 171
QY 754 GTGATGGTCTAAATAGCATGAGTGTACAGCTGGAGCGCCCGGATCCAGATTTTCTTAACC 813
Db 172 LeuThrTrpValAsnSerSerValArgTrpAlaThrArgIleGlnAspMetPheThr 191
QY 814 TTTTGAAGCTCACAGCAATCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTA 873
Db 192 GlyGlyLysLeuLeuAlaLeuSerLeuIleIleGlyValGlyLeuLeuGlnIlePheGln 211
QY 874 GGTCAAACGCAGAACTTTAAA-----GAGCGCTTTTCA---GGAAGAGATTCAAGTATT 924
Db 212 GlyHisPheGluGluLeuArgProSerAsnAlaPheAlaPheTrpMetThrProSerVal 231
QY 925 ACGCGGTTGCCACTGGCTTTTATATGGAATGTATGCATATGCTGGCTGGTTTACCTC 984
Db 232 GlyHisLeuAlaLeuAlaPheLeuGlnGlySerPheAlaPheSerGlyTrpAsnPheLeu 251
QY 985 AACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAACCATTCCTCCCTTGCAATATGTATA 1044
Db 1044

Db 252 AsnTyrValThrGluGluMetValAspAlaArgLysAsnLeuProArgAlaIlePheIle 271
QY 1045 TCCATGGCCATTGTCCACCATTTGGCTATGTGCTGACAAAATGTGGCCTACTTTACGACCAAT 1104
Db 272 SerIleProLeuValThrPheValTyrThrPheThrAsnIleAlaTyrPheThrAlaMet 291
QY 1105 AATGCTGAGGAGCTGCTGCTTTCAAATGAGTGGCAGTGACCTTTTCTGAGCGGCTACTG 1164
Db 292 SerProGlnGluLeuLeuSerSerAsnAlaValAlaValThrPheGlyGluLysLeuLeu 311
QY 1165 GGAATTTCTCATTAGCAGTTCGGATCTTTGTTGCCCTCTCTCGCTGCTCCATGAAC 1224
Db 312 GlyTyrPheSerTrpValMetProValSerValAlaLeuSerThrPheGlyGlyIleAsn 331
QY 1225 GGTGGTGTGTTGCTGTCTCCAGGTTTATTTGTTGCCCTCTCTCGCTGCTCCATGAAC 1284
Db 332 GlyTyrLeuPheThrTyrSerArgLeuCysPheSerGlyAlaArgGluGlyHisLeuPro 351
QY 1285 GAAATCCTCTCCATGATTCTATGTCGCAAGCACACACTCTCTACCAGCTGTTATTTGTTG 1344
Db 352 SerLeuLeuAlaMetIleHisValArgHisCysThrProIleProAlaLeuLeuValCys 371
QY 1345 CACCTTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTC 1404
Db 372 CysGlyAlaThrAlaValIleMetLeuValGlyAspThrTyrThrLeuIleAsnTyrVal 391
QY 1405 AGTTTTGCGAGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATAC 1464
Db 392 SerPheIleAsnTyrLeuCysTyrGlyValThrIleLeuGlyLeuLeuLeuArgTrp 411
QY 1465 AAATGCCAGATATGCATCGTCCTTTTCAAGGTGCCACTGTTTCATCCCAGCTTTGTTTCC 1524
Db 412 ArgArgProAlaLeuHisArgProIleLysValAsnLeuLeuIleProValAlaTyrLeu 431
QY 1525 TTCACATGCCCTCTTCATGTTGCCCTTTTCCCTCTATTCGGACCCATTTAGTACAGGGATT 1584
Db 432 ValPheTrpAlaPheLeuLeuValPheSerPheIleSerGluProMetValCysGlyVal 451
QY 1585 GGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTTATCTCTTTATATATGGACAAG 1644
Db 452 GlyValIleIleIleLeuThrGlyValProIlePhePheLeuGlyValPheTrpArgSer 471
QY 1645 AAACCCAGGTGGTTTAGAATAATGTACAGAGAAAATAACCAACAAATAATACTG 1704
Db 472 LysProLysCysValHisArgLeuThrGluSerMetThrHisTrpGlyGlnGluLeuCys 491
QY 1705 GAAGTTGTA-----CCAGAGAAGATAAGTTATGAACATAATGGACTTGAG 1749
Db 492 PheValValTyrProGlnAspAlaProGluGluGluGlu-----AsnGly----- 506
QY 1750 ATCTTGGCAATCTGCCCA 1767
Db 507 -----ProCysPro 509

Search completed: April 16, 2003, 16:29:30
Job time : 80.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:24:20 ; Search time 70.5 Seconds
(without alignments)
8463.777 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 732
Sequence: 1 ggaggtgaagtgcagag.....ttattataaaaaaaaaa 2239

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues
Word size: 1
Total number of hits satisfying chosen parameters: 1687582
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_091001_21186/app_query.fasta_1.2375
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=olin2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40 cdi
-LIST=45 -DOCLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_1_79@runat_04042003_091001_21186 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	466	63.7	521	22	ABB11957	Human cysteine/Glu
2	424	57.9	424	22	AAM23914	Human EST encoded
3	16	2.2	16	22	AAE13851	Human T cell epit
4	12	1.6	511	20	AA41159	Human amino acid p
5	12	1.6	511	20	AA427076	Human amino acid p
6	12	1.6	511	21	AA487630	Human monp-2 prote
7	10	1.4	76	22	AAM93165	Human digestive sy
8	10	1.4	523	22	AAG63804	Amino acid sequenc
9	10	1.4	530	22	AAG63803	Amino acid sequenc
10	9	1.3	43	22	AAO12595	Human polypeptide
11	9	1.3	62	22	AAO09529	Human polypeptide
12	9	1.2	87	22	ABG19765	Novel human diagno
13	9	1.3	90	22	AAM85927	Novel immune/haema
14	9	1.3	119	22	AAU31764	Novel human secret
15	9	1.2	132	20	AA438429	Human secreted pro
16	9	1.2	149	22	AAO06767	Human polypeptide
17	9	1.2	172	20	AA438417	Human secreted pro
18	9	1.2	188	22	ABG04555	Novel human diagno
19	9	1.2	330	22	AAE03339	Human gene 15 enco
20	9	1.2	330	23	ABG64440	Human albumin fusi
21	8	1.1	20	18	AAW42938	Immunogenic Hepati
22	8	1.1	21	22	ABG69457	Synthetic HAV P2C
23	8	1.1	34	22	AB43020	Peptide #10526 enc
24	8	1.1	34	22	ABB26212	Protein #8211 enco
25	8	1.1	34	22	AAM63925	Human brain expres
26	8	1.1	34	22	AAM76743	Human bone marrow
27	8	1.1	34	22	AAM20989	Peptide #7423 enco
28	8	1.1	34	22	AAM36850	Peptide #10887 enc
29	8	1.1	34	23	ABG40656	Human peptide enco
30	8	1.1	34	23	ABG45922	Human peptide enco
31	8	1.1	43	22	AA64451	Human secreted pro
32	8	1.1	46	22	AAO04640	Human polypeptide
33	8	1.1	47	22	AAE03401	Human gene 9 enco
34	8	1.1	47	23	ABG63681	Human albumin fusi
35	8	1.1	49	22	AB44269	Peptide #11775 enc
36	8	1.1	49	22	ABB27136	Protein #9135 enco
37	8	1.1	49	22	AAM65305	Human brain expres
38	8	1.1	49	22	AAM78004	Human bone marrow
39	8	1.1	49	22	AAM21894	Peptide #8328 enco
40	8	1.1	49	22	AAM38220	Peptide #12257 enc
41	8	1.1	49	23	ABG47017	Human peptide enco
42	8	1.1	54	22	AAO04606	Human polypeptide
43	8	1.1	55	22	ABG00578	Novel human diagno
44	8	1.1	55	22	ABG00581	Novel human diagno
45	8	1.1	58	22	AAM91179	Human immune/haema

ALIGNMENTS

RESULT 1
ABB11957
ID ABB11957 standard; peptide; 521 AA.
XX
AC ABB11957;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human cysteine/Glu transporter homologue, SEQ ID NO:2327.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;

KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer.

XX Homo sapiens.
XX WO200157188-A2.

PN 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US03800.
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.
DR N-PSDB; ABA09201.

XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer --

XX Claim 20; Page 286; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.

CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
XX polypeptide of the invention.

SQ Sequence 521 AA;

Alignment Scores: 0 Length: 521
Pred. No.: 466.00 Matches: 466
Score: 100.00% Conservative: 0
Percent Similarity:

Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 63.66%		Indels: 0
DB: 22		Gaps: 0
US-09-667-170A-440 (1-2239) x ABB11957 (1-521)		
QY 334	GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTAGGGGAGTCTCCATTCATTGCGC	393
Db 56	GlulysValGlnLeuLysArgLysValThrLeuLeuArgGlyValSerIleIleGly	75
QY 394	ACCATCATTCGACGAGGAATCTTCATCTCTCTAAGGGCGTCTCCAGAACACGGGCGAGC	453
Db 76	ThrIleIleGlyAlaGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer	95
QY 454	GTGGGCATGTCTTGACCATCTGGACGGTGTGTGGGGTCTCCGTCCTACTATTGGAGCTTTG	513
Db 96	ValGlyMetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeu	115
QY 514	TCTTATGTGAATGGGAACAACATAAAGAAATCTGGAGGTCAATACACATATATTTG	573
Db 116	SerTyrAlaGluLeuGlyThrThrIleLysLysSerGlyGlyHisTyrThrTyrIleLeu	135
QY 574	GAGTCTTTGGTCCATTACACAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAACGC	633
Db 136	GluValPheGlyProLeuProAlaPheValArgValTrpValGluLeuIleIleArg	155
QY 634	CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTCCTGGAAACCATTTT	693
Db 156	ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPhePhe	175
QY 694	ATTCATGTGAATCCCTGAATTCGATCAAGCTCAAGCTGTGGGCAATAAATGTA	753
Db 176	IleGlnCysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrVal	195
QY 754	GTGATGGTCTAAATAGCATGAGTGTGACGCTGGAGCGCCGGATCCAGATTCTTAAACC	813
Db 196	ValMetValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThr	215
QY 814	TTTTGCAAGCTCACAGCAATCTGTATAATTATAGTCCCTGGAGTTATGCAGCTAATAA	873
Db 216	PheCysLysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLys	235
QY 874	GGTCAACGCGAGAACTTTAAAGACGCCCTTTTTCAGGAAGAGATCAAGTATTACGGGTTG	933
Db 236	GlyGlnThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerIleThrArgLeu	255
QY 934	CCACTGGCTTTTATATGGAATGTATGATATGCTGGTGGTGTTCCTCAACTTGTGT	993
Db 256	ProLeuAlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheVal	275
QY 994	ACTGAAGAAGTAGAAACCCCTGAAAAACCATTCCTCCCTTGCAATATGTATATCCATGGCC	1053
Db 276	ThrGluGluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAla	295
QY 1054	ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG	1113
Db 296	IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrThrIleAsnAlaGlu	315
QY 1114	GAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC	1173
Db 316	GluLeuLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe	335
QY 1174	TCATTAGCAGTCCGATCTTTGTGGCTCTCTCCTGCTCCATGAACGGTGGTGTG	1233
Db 336	SerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGlyVal	355
QY 1234	TTTGTGCTCTCAGGTTATTCTATGTTGGTCTCGAGAGGGTCACCTTCCAGAAATCCTC	1293
Db 356	PheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeu	375
QY 1294	TCCATGATTCATGTCGCGCAAGCACACTCTCTACCAGCTGTATTGTTTGCACCTTTG	1353
Db 376	SerMetIleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeu	395

QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCTCAGTTTGCC 1413
Db |||||||MetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAla 415
QY 1414 AGGTGGCTTTTATTTGGCTGGCAGTGTGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db |||||||LeuPheLeuAlaValAlaGlyLeuLeuTyrLeuArgTyrLysCysPro 435
QY 1474 GATATGCATCTCTCTCAAGGTGGCAGTGTTCATCCAGCTTTTGTCTTCACATGC 1533
Db |||||||AspMetHisArgProPheLysValProLeuPheLeuProAlaLeuPheSerPheThrCys 455
QY 1534 CTCCTTCATGGTGGCTTCCCTCTATTCGGACCCATTTAGTACAGGGATGGCTTCGTC 1593
Db |||||||LeuPheMetValAlaLeuSerLeuTyrSerAspProPheSerThrGlyLeuGlyPheVal 475
QY 1594 ATCCTCTGACTGGAGTCCCTGCGTATTATCTCTTATTTATATGGGACAAAGAACCCAGG 1653
Db |||||||IleThrLeuThrGlyValProAlaTyrTyrLeuPheIleIleTyrAspLysLysProArg 495
QY 1654 TGGTTTAGAATAATGTGAGAGAAAATAACAGAACATTTACAAATACTGGAAGTTGTA 1713
Db |||||||TyrPheArgIleMetSerGluLysIleThrArgThrLeuGlnIleIleLeuGluValVal 515
QY 1714 CCAGAAGAAGATAAGTTA 1731
Db |||||||ProGluGluAspLysLeu 521

RESULT 2
AAM23914
ID AAM23914 standard; Protein; 424 AA.
XX AC AAM23914;
XX DT 12-OCT-2001 (first entry)

DE Human EST encoded protein SEQ ID NO: 1439.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition.

XX Homo sapiens.

XX WO200154477-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02687.

XX 25-JAN-2000; 2000US-0491404.

XX 17-JUL-2000; 2000US-0617746.

XX 03-AUG-2000; 2000US-0631451.

XX 15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;

XX Cao Y, Drmanac RA, Zhang J, Werhman T;

XX WPI; 2001-476164/51.

XX N-PSDB; AAH98573.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use -

XX Claim 20; Page 1001-1002; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
XX proteins from a variety of organisms, including human, dog, cat, horse,
XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea

CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.
XX

SQ Sequence 424 AA;

Alignment Scores:

Pred. No.: 0 Length: 424
Score: 424.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.92% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAM23914 (1-424)

QY 460 ATGTCTCTGACCATCTGGACGGTGTGGGGTCTGTACATATTGGAGCTTTGTCTTAT 519

Db 1 MetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeuSerTyr 20

QY 520 GCTCAATTGGGAACAACCTATAAAGAATCTGGAGGTCAATACACATATATTTGGAGTC 579

Db 21 AlaGluLeuGlyThrThrIleLysLysSerGlyGlyHisTyrThrIleLeuGluVal 40

QY 580 TTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATACGCCCTGCA 639

Db 41 PheGlyProLeuProAlaPheValArgValTrpValGluLeuLeuIleIleArgProAla 60

QY 640 GCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTGGAACCATTTTATTCAA 699

Db 61 AlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPhePheIleGln 80

QY 700 TGTGAATCCCTGAACCTTGGCATCAAGCTCATTAACAGTGTGGGCATTAAGTGTG 759

Db 81 CysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrValValMet 100

QY 760 GTCCTAAATAGCATGAGTGTGACGCTGGAGCGCGGATCCAGATTTTCTTAACCTTTGC 819

Db 101 ValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThrPheCys 120

QY 820 AAGCTCACAGCAATCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAA 879

Db 121 LysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLysGlyGln 140

QY 880 ACGCAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTG 939

Db 141 ThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerSerIleThrArgLeuProLeu 160

QY 940 GCTTTTATTATGGAATGTATGCATATGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGT 999

Db 161 AlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTyrPheTyrLeuAsnPheValThrGlu 180

QY 1000 GAAGTAGAAAACCCCTGAAAAACCAATCCCTTGCATATATCCATGGCCATTGTC 1059

Db 181 GluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAlaIleVal 200

QY 1060 ACCATTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAGAGCTG 1119

Db 201 ThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrIleAsnAlaGluGluLeu 220

QY 1120 CTGCTTTCAAATGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGT 1179

Db 221 LeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPheSerLeu 240

QY 1180 GCAGTTCGATCTTTGTTGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239

Db 241 AlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGlyValPheAla 260

QY 1240 GTCTCCAGGTTATTCTATGTGGCTCTCGAGAGGTCACCTTCCAGAAATCTCTCCATG 1299

Db 261 ValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeuSerMet 280

XX The invention relates to isolated lung tumour-specific proteins and
CC their corresponding cDNA molecules. Lung tumour-specific proteins and
CC their antigen-presenting cells are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, and for inhibiting
CC the development of cancer. The invention also relates to a composition
CC useful for stimulating an immune response, and for treating cancer. The
CC lung tumour specific oligonucleotide is useful in gene therapy and for
CC diagnosis, detection and treatment of lung cancer. The present sequence
CC is human T cell epitope related to lung tumour-specific protein.
XX
SQ Sequence 16 AA;

Alignment Scores:
Pred. No.: 8.19e-07 Length: 16
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.19% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAEL13851 (1-16)

QY 104 TTCCAGCCCAATTGTGGCATAGATTTTATCATATCTCTGGATTTTGG 151
|||||
Db 1 PheGlnAlaAsnCysGlylleAspPheIlellePheTrpIlePheTrp 16

RESULT 4
AAEL13851
ID AAY41159 standard; Protein; 511 AA.
XX
AC AAY41159;
XX
DT 24-JAN-2000 (first entry)
XX
DE Human amino acid permease homolog (AAPH).
XX
KW Amino acid permease homolog; AAPH; cancer; inflammatory disorder; AIDS;
KW autoimmune disorder; anemia; asthma; atherosclerosis; diabetes mellitus;
KW multiple sclerosis; osteoarthritis; osteoporosis; rheumatoid arthritis;
KW ulcerative colitis; infection; cell proliferation disorder; human.
XX
OS Homo sapiens.
XX
PN US5981242-A.
XX
PD 09-NOV-1999.
XX
PF 03-FEB-1999; 99US-0243920.
XX
PR 06-MAY-1998; 98US-0073362.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Yue H, Corley NC, Hillman JL;
XX
DR WPI; 1999-633325/54.
DR N-PSDB; AAZ23240.
XX
PT New human amino acid permease homolog, useful in the diagnosis,
PT treatment and prevention of cancer, inflammatory/autoimmune disorders
PT and cell proliferation disorders -
XX
PS Claim 1; Fig 1A-F; 32pp; English.
XX
CC This represents a human amino acid permease homolog (AAPH). The AAPH
CC nucleic acid, protein, AAPH (ant)agonists are useful in the diagnosis,
CC treatment and prevention of cancer such as adenocarcinoma, leukemia,
CC lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the
CC adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall
CC bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung,
CC muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands,
CC skin, spleen, testis, thymus, thyroid or uterus; inflammatory/autoimmune

QY 1300 ATTCATGTCGCGAAGCACACTCTCTACAGCTGTATTGTTTGCACCTTTGACAATG 1359
|||||
Db 281 IleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeuThrMet 300

QY 1360 ATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGTCCAGGTGG 1419
|||||
Db 301 IleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAlaArgTrp 320

QY 1420 CTTTATTTATGGGCTGGCAGTTCGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATG 1479
|||||
Db 321 LeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysProAspMet 340

QY 1480 CATCGTCCCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTTCACATGCTCTTC 1539
|||||
Db 341 HisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCysLeuPhe 360

QY 1540 ATGGTTGCCCTTCCCTCTATTCCGACCCCAATTTAGTACAGGATTTGGCTTCGTCATCACT 1599
|||||
Db 361 MetValAlaLeuSerLeuTyrSerAspPropheSerThrGlyIleGlyPheValIleThr 380

QY 1600 CTGACTGGAGTCCCTCGCTATTATCTCTTATTTATATGTTGGACAAACCCAGGTGTTT 1659
|||||
Db 381 LeuThrGlyValProAlaTyrTyrLeuPheIleIleTrpAspLysLysProArgTrpPhe 400

QY 1660 AGAATAATGTGAGAGAAATAACAGAACATTACAAATAATACTGGAAGTTGTACCAGAA 1719
|||||
Db 401 ArgIleMetSerGluLysIleThrArgThrLeuGlnIleIleLeuGluValValProGlu 420

QY 1720 GAAGATAAGTTA 1731
|||||
Db 421 GluAspLysLeu 424

RESULT 3
AAEL13851
ID AAE13851 standard; peptide; 16 AA.
XX
AC AAE13851;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human T cell epitope related to lung tumour-specific protein.
XX
KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer;
KW T cell epitope.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US09991.
XX
PR 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-234517P.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
DR WPI; 2001-639201/73.
XX
PT New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer -
XX
PS Claim 2; Page 378; 378pp; English.

CC disorders such as AIDS, Addison's disease, adult respiratory distress
CC syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia,
CC asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune
CC thyroiditis, bronchitis, cholecystitis, contact dermatitis, diabetes
CC mellitus, Crohn's disease, atopic dermatitis, dermatomyositis,
CC emphysema, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid
CC arthritis, ulcerative colitis, viral, bacterial, fungal, parasitic,
CC protozoal infections; and cell proliferation disorders such as actinic,
CC keratosis, cirrhosis, arteriosclerosis, myelofibrosis, and primary
CC thrombocytopenia.
XX
SQ Sequence 511 AA;

Alignment Scores:
Pred. No.: 0.00575 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 20 Gaps: 0

US-09-667-170A-440 (1-2239) x AAY41159 (1-511)

QY 517 TATGCTGAATTGGGAACAACACTATAAAGAAATCTGGA 552
|||||
Db 90 TyrAlaGluLeuGlyThrThrIleLysLysSerGly 101

RESULT 5
AAY27076

ID AAY27076 standard; Protein; 511 AA.
XX
AC AAY27076;

DT 18-OCT-1999 (first entry)

DE Human amino acid permease homologue (AAPH).

XX Amino acid permease homologue; AAPH; cancer; inflammatory disorder;
KW autoimmune disorder; cell proliferation disorder; adenocarcinoma;
KW AIDS; Crohn's disease; rheumatoid arthritis; arteriosclerosis; human;
KW psoriasis; hepatitis; gene therapy.

OS Homo sapiens.

XX US5942399-A.

PN 24-AUG-1999.

PD 06-MAY-1998; 98US-0073362.

PF 06-MAY-1998; 98US-0073362.

PR (INCY-) INCYTE PHARM INC.

XX Corley NC, Hillman JL, Yue H;

PI WPI; 1999-493508/41.

DR N-PSDB; AAX86999.

XX New human amino acid permease homologue (AAPH) polynucleotide and its
PT complement useful for prevention and treatment of cancer, autoimmune
PT and cell proliferation disorders

XX Claim 1; Fig 1A-F; 33pp; English.

CC This sequence represents a human amino acid permease homologue (AAPH).
CC The protein can be expressed by standard recombinant methodology. AAPH
CC polynucleotides and polypeptides are useful for diagnosis, treatment and
CC prevention of cancer e.g. adenocarcinoma, lung cancer, ovarian cancer;
CC inflammatory/autoimmune disorders e.g. AIDS, Crohn's disease, rheumatoid
CC arthritis; and cell proliferation disorders e.g. arteriosclerosis,
CC psoriasis, hepatitis. AAPH polynucleotides are useful in gene therapy,
CC in PCR-based methods for detecting upstream sequences, and also for

CC generating hybridisation probes for mapping the naturally occurring
CC genomic sequence.
XX
SQ Sequence 511 AA;

Alignment Scores:
Pred. No.: 0.00575 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 20 Gaps: 0

US-09-667-170A-440 (1-2239) x AAY27076 (1-511)

QY 517 TATGCTGAATTGGGAACAACACTATAAAGAAATCTGGA 552
|||||
Db 90 TyrAlaGluLeuGlyThrThrIleLysLysSerGly 101

RESULT 6
AAY87630

ID AAY87630 standard; Protein; 511 AA.

XX AAY87630;

DT 01-AUG-2000 (first entry)

XX Human monp-2 protein.

DE Human; monp-2; cell proliferation; drug development; immune disease;
XX drug development; cancer.

OS Homo sapiens.

XX WO200017238-A1.

PN 30-MAR-2000.

PD 22-SEP-1999; 99WO-JP05205.

PF 24-SEP-1998; 98JP-0270356.

PR (TAIS) TAISHO PHARM CO LTD.

XX Yoshimoto M, Takayama K;

PI WPI; 2000-283547/24.

DR N-PSDB; AAA11954.

XX Novel gene monp-2 encoding monocyte-specific protein with cell
PT proliferation promoting effect, useful in studying its expression and
PT immune function, and in developing drugs for e.g. immune diseases and
PT cancers

XX Claim 1a; Page 22-25; 34pp; Japanese.

XX This invention describes a novel protein MONP-2 (I) and its encoding
CC gene monp-2 (II) with a cell proliferation promoting effect isolated
CC from a human peripheral blood. The new protein (I) and gene are useful
CC in studying the expression abnormality and dysfunction of (I) as well
CC as maintenance of immune function and development of drugs for e.g.
CC immune diseases and cancers. This represents encodes the MONP-2 protein
CC described in the method of the invention.

XX SQ Sequence 511 AA;

Alignment Scores:
Pred. No.: 0.00575 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 21 Gaps: 0

```

US-09-667-170A-440 (1-2239) x AAY87630 (1-511)

QY 517 TATGCTGAATTGGGAACAACACTATAAGAAATCTGGA 552
    |||||||||||||||||||||||||||||||||||
Db 90 TyrAlaGluLeuGlyThrThrIleLysLysSerGly 101

RESULT 7
AAM93165
ID AAM93165 standard; Protein; 76 AA.
XX
AC AAM93165;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen SEQ ID NO: 2514.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

```


XX The present sequence represents a sodium independent small neutral amino
CC acid transporter polypeptide. The polypeptides and polynucleotides
CC are used for controlling cell resistance to oxidizing stress, NMDA type
CC glutamate receptor activity, cell growth and for screening for
CC substances that have these activities.

SQ Sequence 530 AA;

Alignment Scores:
Pred. No.: 0.618 Length: 530
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.37% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAG63803 (1-530)

QY 409 GGAATCTTCATCTCTCCTAAGGGCGTGCTC 438
|||||
Db 63 GlyIlePheIleSerProLysGlyValLeu 72

RESULT 10
AAO12595
ID AAO12595 standard; Protein; 43 AA.

XX AAO12595;

AC AAO12595;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 26487.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI92526.

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 20; SEQ ID NO 26487; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 43 AA;

Alignment Scores:
Pred. No.: 9.27 Length: 43
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAO12595 (1-43)

QY 54 GACAGAGTCTCACTGTCCACCCAGGCTG 28
|||||

Db 1 AspArgValSerLeuSerProArgLeu 9

RESULT 11

AAO09529

ID AAO09529 standard; Protein; 62 AA.

XX AAO09529;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 23421.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI89460.

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 20; SEQ ID NO 23421; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 62 AA;
Alignment Scores:
Pred. No.: 8.78 Length: 62
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAO09529 (1-62)

QY 29 TGGCATGATCTGCTCACTTCAACCT 3

Db 7 TrpHisAspLeuCysSerLeuGlnPro 15

RESULT 12

ABG19765
ID ABG19765 standard; Protein; 87 AA.

XX AC ABG19765;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19756.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS83952.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX PS Claim 20; SEQ ID No 50124; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 87 AA;

Alignment Scores:
Pred. No.: 8.36 Length: 87
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x ABG19765 (1-87)

QY 72 AAAAAGAAAGAAAGAAAGAGAGAGA 98

Db 18 LysLysLysGluArgLysArgGluArg 26

RESULT 13

AAM85927

ID AAM85927 standard; Protein; 90 AA.

XX AC AAM85927;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen SEQ ID NO:13520.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01354.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 11-JUL-2000; 2000US-0217496.

XX PR 14-JUL-2000; 2000US-0218290.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 26-JUL-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225266.

XX PR 14-AUG-2000; 2000US-0225267.

XX PR 14-AUG-2000; 2000US-0225268.

XX PR 14-AUG-2000; 2000US-0225270.

XX PR 14-AUG-2000; 2000US-0225447.

XX PR 14-AUG-2000; 2000US-0225757.

XX PR 14-AUG-2000; 2000US-0225758.

XX PR 14-AUG-2000; 2000US-0225759.

XX PR 18-AUG-2000; 2000US-0226279.

XX PR 22-AUG-2000; 2000US-0226681.

PR	22-AUG-2000;	2000US-0226868.	PR	17-NOV-2000;	2000US-0249208.
PR	22-AUG-2000;	2000US-0227182.	PR	17-NOV-2000;	2000US-0249209.
PR	23-AUG-2000;	2000US-0227009.	PR	17-NOV-2000;	2000US-0249210.
PR	30-AUG-2000;	2000US-0228924.	PR	17-NOV-2000;	2000US-0249211.
PR	01-SEP-2000;	2000US-0229287.	PR	17-NOV-2000;	2000US-0249212.
PR	01-SEP-2000;	2000US-0229343.	PR	17-NOV-2000;	2000US-0249213.
PR	01-SEP-2000;	2000US-0229344.	PR	17-NOV-2000;	2000US-0249214.
PR	01-SEP-2000;	2000US-0229345.	PR	17-NOV-2000;	2000US-0249215.
PR	05-SEP-2000;	2000US-0229509.	PR	17-NOV-2000;	2000US-0249216.
PR	05-SEP-2000;	2000US-0229513.	PR	17-NOV-2000;	2000US-0249217.
PR	06-SEP-2000;	2000US-0230437.	PR	17-NOV-2000;	2000US-0249218.
PR	06-SEP-2000;	2000US-0230438.	PR	17-NOV-2000;	2000US-0249244.
PR	08-SEP-2000;	2000US-0231242.	PR	17-NOV-2000;	2000US-0249245.
PR	08-SEP-2000;	2000US-0231243.	PR	17-NOV-2000;	2000US-0249264.
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	2000US-0231413.	PR	17-NOV-2000;	2000US-0249297.
PR	08-SEP-2000;	2000US-0231414.	PR	17-NOV-2000;	2000US-0249299.
PR	08-SEP-2000;	2000US-0232080.	PR	17-NOV-2000;	2000US-0249300.
PR	08-SEP-2000;	2000US-0232081.	PR	01-DEC-2000;	2000US-0250160.
PR	12-SEP-2000;	2000US-0231968.	PR	01-DEC-2000;	2000US-0250391.
PR	14-SEP-2000;	2000US-0232397.	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	2000US-0232398.	PR	05-DEC-2000;	2000US-0251988.
PR	14-SEP-2000;	2000US-0232399.	PR	05-DEC-2000;	2000US-0256719.
PR	14-SEP-2000;	2000US-0232400.	PR	06-DEC-2000;	2000US-0251479.
PR	14-SEP-2000;	2000US-0232401.	PR	08-DEC-2000;	2000US-0251856.
PR	14-SEP-2000;	2000US-0233063.	PR	08-DEC-2000;	2000US-0251868.
PR	14-SEP-2000;	2000US-0233064.	PR	08-DEC-2000;	2000US-0251869.
PR	14-SEP-2000;	2000US-0233065.	PR	08-DEC-2000;	2000US-0251989.
PR	21-SEP-2000;	2000US-0234223.	PR	08-DEC-2000;	2000US-0251990.
PR	21-SEP-2000;	2000US-0234274.	PR	11-DEC-2000;	2000US-0254097.
PR	25-SEP-2000;	2000US-0234997.	PR	05-JAN-2001;	2001US-0259678.
PR	25-SEP-2000;	2000US-0234998.	XX		
PR	26-SEP-2000;	2000US-0235484.	PA		
PR	27-SEP-2000;	2000US-0235834.	XX		
PR	27-SEP-2000;	2000US-0235836.	PI		
PR	29-SEP-2000;	2000US-0236327.	XX		
PR	29-SEP-2000;	2000US-0236367.	DR		
PR	29-SEP-2000;	2000US-0236368.	DR		
PR	29-SEP-2000;	2000US-0236369.	XX		
PR	29-SEP-2000;	2000US-0236370.	PT		
PR	02-OCT-2000;	2000US-0236802.	PT		
PR	02-OCT-2000;	2000US-0237037.	XX		
PR	02-OCT-2000;	2000US-0237038.	PS		
PR	02-OCT-2000;	2000US-0237039.	XX		
PR	13-OCT-2000;	2000US-0237040.	CC		
PR	13-OCT-2000;	2000US-0239935.	CC		
PR	13-OCT-2000;	2000US-0239937.	CC		
PR	20-OCT-2000;	2000US-0240960.	CC		
PR	20-OCT-2000;	2000US-0241221.	CC		
PR	20-OCT-2000;	2000US-0241785.	CC		
PR	20-OCT-2000;	2000US-0241786.	CC		
PR	20-OCT-2000;	2000US-0241787.	CC		
PR	20-OCT-2000;	2000US-0241808.	CC		
PR	20-OCT-2000;	2000US-0241809.	CC		
PR	20-OCT-2000;	2000US-0241826.	CC		
PR	01-NOV-2000;	2000US-0244617.	CC		
PR	08-NOV-2000;	2000US-0246474.	CC		
PR	08-NOV-2000;	2000US-0246475.	CC		
PR	08-NOV-2000;	2000US-0246476.	CC		
PR	08-NOV-2000;	2000US-0246477.	CC		
PR	08-NOV-2000;	2000US-0246478.	CC		
PR	08-NOV-2000;	2000US-0246523.	XX		
PR	08-NOV-2000;	2000US-0246524.	SQ		
PR	08-NOV-2000;	2000US-0246525.			
PR	08-NOV-2000;	2000US-0246526.			
PR	08-NOV-2000;	2000US-0246527.			
PR	08-NOV-2000;	2000US-0246528.			
PR	08-NOV-2000;	2000US-0246532.			
PR	08-NOV-2000;	2000US-0246609.			
PR	08-NOV-2000;	2000US-0246610.			
PR	08-NOV-2000;	2000US-0246611.			
PR	08-NOV-2000;	2000US-0246613.			
PR	17-NOV-2000;	2000US-0249207.			

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

N-PSDB; AAK58708.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Claim 11; SEQ ID NO 13520; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.

Alignment Scores:			
Pred. No.:	8.32	Length:	90
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.26%	Indels:	0
DB:	22	Gaps:	0

US-09-667-170A-440 (1-2239) x AAM85927 (1-90)

QY 29 TGGCATGATCTCTGCTCACTTCAACCT 3
Db 49 TrpHisAspLeuCysSerLeuGlnPro 57

RESULT 14

AAU31764
ID AAU31764 standard; Protein; 119 AA.
XX
AC AAU31764;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #2255.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.
XX
XX WO200179449-A2.
PN
PD 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US08656.
PF
XX 18-APR-2000; 2000US-0552929.
PR
XX 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
DR
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 505; 765pp; English.
XX

The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid sequences of novel human secreted proteins of the invention.

SQ Sequence 119 AA;

Alignment Scores:

Pred. No.: 7.99 Length: 119
Score: 9.00 Matches: 9
Best Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAU31764 (1-119)

QY 29 TGGCATGATCTCTGCTCACTTCAACCT 3

Db 17 TrpHisAspLeuCysSerLeuGlnPro 25

RESULT 15

AAU38429
ID AAY38429 standard; Protein; 132 AA.
XX
AC AAY38429;
XX
DT 30-SEP-1999 (first entry)
XX
DE Human secreted protein encoded by gene No. 32.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.
XX
XX WO9935158-A1.
PN
PD 15-JUL-1999.
XX
PF 06-JAN-1999; 99WO-US00108.
XX
PR 07-JAN-1998; 98US-0070704.
PR 07-JAN-1998; 98US-0070657.
PR 07-JAN-1998; 98US-0070658.
PR 07-JAN-1998; 98US-0070692.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
XX
DR WPI; 1999-444190/37.
DR
XX N-PSDB; AAU06262.
XX

New isolated human genes and the secreted polypeptides they encode

Claim 11; Page 202-203; 227pp; English.

This sequence represents a secreted human protein encoded by the gene clone detailed in the descriptor line. The gene can be used to generate fusion proteins by linking to the gene to a human immunoglobulin Fc portion (e.g. AAU06210) for increasing the stability of the fused protein as compared to the human protein only.

The invention relates to 36 novel genes and their fragments (nucleic acid sequences: AAU06219-206263; amino acid sequences AAY38386-Y38498) which are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Also, pathological polypeptides can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the 36 polynucleotides, based on which tissues they are most highly expressed in (see AAU06219 for described uses).

SQ Sequence 132 AA;

Alignment Scores:

Pred. No.: 7.87 Length: 132
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 20 Gaps: 0

US-09-667-170A-440 (1-2239) x AAY38429 (1-132)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:29:36 ; Search time 43.5 Seconds
(without alignments)
9896.320 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 732
Sequence: 1 ggaggttgagtgagcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_091002_21212/app_query.fasta_1.2375
-DB=PIR_73 -QFMT=fastan -SUFFIX=oligo -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_1_57@runat_04042003_091002_21212 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_73:*

1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	8	1.1	46	1 VCBPPF	coat protein - pha
2	8	1.1	67	2 A70375	ribosomal protein
3	8	1.1	82	2 S15143	major coat protein
4	8	1.1	82	2 C83556	coat protein B of
5	8	1.1	100	2 H95283	conserved hypothet
6	8	1.1	128	2 S75737	hypothetical prote
7	8	1.1	159	2 T05279	transcription fact
8	8	1.1	184	2 H96718	AthVA22c, 50565-49
9	8	1.1	260	1 S11984	fapR protein - Esc
10	8	1.1	287	2 H75098	hypothetical prote
11	8	1.1	348	2 E82933	type I restriction
12	8	1.1	353	2 A42700	replication factor
13	8	1.1	359	2 I50704	replication factor
14	8	1.1	361	2 T51078	related to homeoti

15	8	1.1	366	2	T27257	hypothetical prote
16	8	1.1	399	2	T23470	hypothetical prote
17	8	1.1	404	2	F86445	unknown protein [i
18	8	1.1	419	1	S70617	homeotic protein g
19	8	1.1	464	2	T28818	hypothetical prote
20	8	1.1	483	2	T31138	hypothetical prote
21	8	1.1	485	2	T27147	hypothetical prote
22	8	1.1	486	2	G86320	F6A14.19 protein -
23	8	1.1	487	2	A71407	probable Ste20-lik
24	8	1.1	499	2	G86538	leucyl aminopeptid
25	8	1.1	499	2	G72083	leucine aminopepti
26	8	1.1	505	2	AC3043	aldehyde dehydroge
27	8	1.1	507	2	JG0165	LAT1 protein - hum
28	8	1.1	511	2	A98243	probable aldehyde
29	8	1.1	537	2	T27612	hypothetical prote
30	8	1.1	537	2	T04822	hypothetical prote
31	8	1.1	541	2	T31554	hypothetical prote
32	8	1.1	546	2	T37665	probable t-complex
33	8	1.1	551	2	F69282	medium-chain acyl-
34	8	1.1	554	2	A71729	DNA repair protein
35	8	1.1	562	2	T16854	hypothetical prote
36	8	1.1	582	2	T23624	hypothetical prote
37	8	1.1	617	2	T23623	hypothetical prote
38	8	1.1	662	2	T32821	hypothetical prote
39	8	1.1	679	2	S43129	penicillin-binding
40	8	1.1	725	2	A47168	cardiac morphogene
41	8	1.1	731	2	C85060	hypothetical prote
42	8	1.1	746	2	S31816	genome polyprotein
43	8	1.1	760	2	T06291	extensin homolog T
44	8	1.1	786	2	I49274	mammary gland fact
45	8	1.1	793	2	S54772	mammary gland fact

ALIGNMENTS

RESULT 1

VCBPPF

coat protein - phage Pf1

C;Species: phage Pf1

C;Date: 24-Sep-1981 #sequence_revision 24-Sep-1981 #text_change 31-Dec-1993

C;Accession: A04229

R;Nakashima, Y.; Wiseman, R.L.; Konigsberg, W.; Marvin, D.A.

Nature 253, 68-71, 1975

A;Title: Primary structure and sidechain interactions of PFL filamentous bacterial vi

A;Reference number: A04229; MUID:75082355; PMID:1110754

A;Accession: A04229

A;Molecule type: protein

A;Residues: 1-46 <NAK>

C;Comment: Pf1 is a filamentous bacteriophage with single-stranded DNA.

C;Superfamily: filamentous phage coat protein B

Alignment Scores:

Pred. No.:	55	Length:	46
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	1	Gaps:	0

US-09-667-170A-440 (1-2239) x VCBPPF (1-46)

Qy 1432 CTGGCAGTTGCTGGGCTGATTAT 1455

|||||

Db 33 LeuAlaValAlaGlyLeuIleTyr 40

RESULT 2

A70375

ribosomal protein S21 - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 13-Aug-1999

C;Accession: A70375

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

us-09-667-170a-440.olin2p.rpr

Thu Apr 17 09:48:32 2003

Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: A70375
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-67 <AQF>
A;Cross-references: GB:AE000711; NID:g2983401; PIDN:AAC06990.1; PID:g2983406; GB:AE000069
A;Experimental source: strain VF5
C;Genetics:
A;Gene: rpsU
C;Superfamily: Escherichia coli ribosomal protein S21

Alignment Scores: 52.9 Length: 67
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.09% Gaps: 0
DB: 2

US-09-667-170A-440 (1-2239) x A70375 (1-67)

QY 77 GAACAAAGAAAAGAGAGAGAGG 100
|||||
Db 43 GluArgLysLysLysArgGluArg 50

RESULT 3
S15143
Major coat protein precursor - phage Pf1
N;Alternate names: gene 8 protein
C;Species: phage pf1
C;Date: 18-Feb-1994 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999
C;Accession: S15143; S20699
R;Hill, D.F.; Short, N.J.; Perham, R.N.; Petersen, G.B.
J. Mol. Biol. 218, 349-364, 1991
A;Title: DNA sequence of the filamentous bacteriophage Pf1.
A;Reference number: S15140; MUID:91186399; PMID:2010913
A;Accession: S15143
A;Molecule type: DNA
A;Residues: 1-82 <JMO>
A;Cross-references: EMBL:X52107; NID:g14829; PIDN:CAA36331.1; PID:g14833
A;Experimental source: ATCC 25102-B1
C;Superfamily: filamentous phage coat protein B
F;1-36/Domain: signal sequence #status predicted <SIG>
F;37-82/Product: major coat protein #status predicted <MAT>

Alignment Scores: 51.8 Length: 82
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.09% Gaps: 0
DB: 2

US-09-667-170A-440 (1-2239) x S15143 (1-82)

QY 1432 CTGGCAGTGTGGCTGATTAT 1455
|||||
Db 69 LeuAlaValAlaGlyLeuIleTyr 76

RESULT 4
C83556
coat protein B of bacteriophage (Pf1) PA0723 [imported] - Pseudomonas aeruginosa (strain
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: C83556
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: C83556
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-82 <STO>
A;Cross-references: GB:AE004507; GB:AE004091; NID:g9946596; PIDN:AAG04112.1; GSPDB:GN
A;Experimental source: strain PA01
C;Genetics:
A;Gene: coaB; PA0723
C;Superfamily: filamentous phage coat protein B

Alignment Scores: 51.8 Length: 82
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.09% Gaps: 0
DB: 2

US-09-667-170A-440 (1-2239) x C83556 (1-82)

QY 1432 CTGGCAGTGTGGCTGATTAT 1455
|||||
Db 69 LeuAlaValAlaGlyLeuIleTyr 76

RESULT 5
H95283
conserved hypothetical protein Sma0325 [imported] - Sinorhizobium meliloti (strain 10
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: H95283
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: H95283
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-100 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK64834.1; PID:g14523247; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma0325
A;Genome: plasmid

Alignment Scores: 50.8 Length: 100
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.09% Gaps: 0
DB: 2

US-09-667-170A-440 (1-2239) x H95283 (1-100)

QY 1075 CTGACAAATGTGGCTACTTTACG 1098
|||||
Db 4 LeuThrAsnValAlaTyrPheThr 11

RESULT 6
S75737
hypothetical protein slr0921 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S75737
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75737
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-128 <KAN>
A;Cross-references: EMBL:D64003; GB:AB001339; NID:g1001200; PIDN:BAA10472.1; PID:g100123
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: *Synechocystis* hypothetical protein slr0921

Alignment Scores:
Pred. No.: 49.5 Length: 128
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x S75737 (1-128)

QY 349 TCAGCTGCACCTTCTCTCTCTCTCTCTG 326

Db 89 SerAlaAlaLeuSerLeuPheLeu 96

RESULT 7
T05279

transcription factor ATB2 - *Arabidopsis thaliana*
N;Alternate names: protein T4L20.170

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000

C;Accession: T05279

R;Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, September 1998

A;Reference number: Z15406

A;Accession: T05279

A;Molecule type: DNA

A;Residues: 1-159 <BEV>

A;Cross-references: EMBL:AL023094

A;Experimental source: cultivar Columbia; BAC clone T4L20

C;Genetics:

A;Gene: ATB2

A;Map position: 4

A;Note: T4L20.170

C;Superfamily: tobacco DNA-binding protein tbz17; fos/jun DNA-binding domain homology

Alignment Scores:

Pred. No.: 48.4 Length: 159

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.12% Indels: 0

DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x T05279 (1-159)

QY 1724 TCTTCTTCTGGTACACCTTCCAGT 1701

Db 4 SerSerSerGlyThrThrSerSer 11

RESULT 8

H96718

ATHVA22c, 50565-49239 [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: H96718

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96718
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-184 <STO>
A;Cross-references: GB:AE005173; NID:g6665546; PIDN:AAF22915.1; GSPDB:GN00141
C;Genetics:

A;Gene: T6C23.10

A;Map position: 1

Alignment Scores:

Pred. No.: 47.7 Length: 184

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x H96718 (1-184)

QY 963 ATATGCTGGCTGGTTTACCTCAA 986

Db 85 IleCysTrpLeuValLeuProGln 92

RESULT 9

S11984

fapR protein - *Escherichia coli*

C;Species: *Escherichia coli*

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: S11984

R;Klaassen, P.; de Graaf, F.K.

Mol. Microbiol. 4, 1779-1783, 1990

A;Title: Characterization of FapR, a positive regulator of expression of the 987P op

A;Reference number: S11984; MUID:91171879; PMID:2077360

A;Accession: S11984

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-260 <KLA>

A;Cross-references: EMBL:X53494

A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in lacking 96

C;Genetics:

A;Gene: fapR

C;Superfamily: fapR protein

C;Keywords: DNA binding; transcription regulation

Alignment Scores:

Pred. No.: 46 Length: 260

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x S11984 (1-260)

QY 1135 GTGGCAGTGACCTTTCTGAGCGG 1158

Db 148 ValAlaValThrPheSerGluArg 155

RESULT 10

H75098

hypothetical protein PAB0790 - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

us-09-667-170a-440.olin2p.rpr

Thu Apr 17 09:48:32 2003

C;Accession: H75098
 R;anonymouse, Genoscope
 submitted to the EMBL Data Library, July 1999
 A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
 A;Reference number: A75001
 A;Accession: H75098
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-287 <KAW>
 A;Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB50093.1; PID:e151599
 A;Experimental source: strain Orsay
 C;Genetics:
 A;Gene: PAB0790

Alignment Scores:
 Pred. No.: 45.6 Length: 287
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x H75098 (1-287)

QY 1369 TTCTCTGGAGACCTCGACAGTCTT 1392

Db 53 PheSerGlyAspLeuAspSerLeu 60

RESULT 11

E82933
 type I restriction enzyme M protein, truncated homolog UU098 [imported] - Ureaplasma ure
 C;Species: Ureaplasma urealyticum
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Aug-2002
 C;Accession: E82933
 R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to Genbank, February 2000
 A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A;Reference number: A82870
 A;Accession: E82933
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-348 <GLA>
 A;Cross-references: GB:AE002110; GB:AF222894; NID:g6899051; PIDN:AAF30504.1; GSPDB:GN001
 A;Experimental source: serovar 3; biovar 1
 C;Genetics:
 A;Gene: hsdM-1; UU098
 A;Genetic code: SGC3
 C;Superfamily: type I site-specific deoxyribonuclease chain hsdM (associate member)

Alignment Scores:
 Pred. NO.: 44.7 Length: 348
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x E82933 (1-348)

QY 1733 CATAACTTATCTTCTTCTGTACA 1710

Db 173 HisAsnLeuSerSerSerGlyThr 180

RESULT 12

A42700
 replication factor C - human
 C;Species: Homo sapiens (man)
 C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Aug-1998
 C;Accession: A42700
 R;Chen, M.; Pan, Z.Q.; Hurwitz, J. 2516-2520, 1992
 Proc. Natl. Acad. Sci. U.S.A. 89, 2516-2520, 1992
 A;Title: Sequence and expression in Escherichia coli of the 40-kDa subunit of activator
 A;Reference number: A42700; MUID:92212860; PMID:1313560

A;Accession: A42700
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-353 <CHE>
 A;Cross-references: GB:M87338
 C;Genetics:
 A;Gene: GDB:RFC2; A1
 A;Cross-references: GDB:283460
 A;Map position: 7q11.23-7q11.23
 C;Superfamily: phage T4 DNA polymerase accessory protein 44

Alignment Scores:
 Pred. NO.: 44.6 Length: 353
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x A42700 (1-353)

QY 1761 CTGCCCAAGGGGAGACACAAAATA 1784

Db 129 LeuProLysGlyArgHisLysIle 136

RESULT 13

I50704
 replication factor C/activator 1 subunit - chicken
 C;Species: Gallus gallus (chicken)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Sep-1999
 C;Accession: I50704
 R;Oberholtzer, J.; Cohen, E.L.; Davis, J.G.
 DNA Cell Biol. 13, 857-863, 1994
 A;Title: Molecular cloning of a chick cochlea cDNA encoding a subunit of DNA replicat
 A;Reference number: I50704; MUID:94347215; PMID:8068208
 A;Accession: I50704
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-359 <OBE>
 A;Cross-references: EMBL:U12438; NID:g527668; PIDN:AAA20552.1; PID:g527669
 C;Superfamily: phage T4 DNA polymerase accessory protein 44

Alignment Scores:
 Pred. NO.: 44.5 Length: 359
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x I50704 (1-359)

QY 1761 CTGCCCAAGGGGAGACACAAAATA 1784

Db 135 LeuProLysGlyArgHisLysIle 142

RESULT 14

T51078
 related to homeotic protein Antennapedia [imported] - Neurospora crassa
 N;Alternate names: protein B2A19.120
 C;Species: Neurospora crassa
 C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C;Accession: T51078
 R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
 submitted to the Protein Sequence Database, July 2000
 A;Reference number: Z25286
 A;Accession: T51078
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-361 <SCH>
 A;Cross-references: EMBL:AL390092; GSPDB:GN00116; NCSP:B2A19.120
 A;Experimental source: BAC clone B2A19; strain OR74A
 C;Genetics:

A;Gene: NCSP:B2A19.120
A;Map position: 6
A;Introns: 282/2

Alignment Scores:
Pred. No.: 44.5 Length: 361
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x T51078 (1-361)

QY 148 TTGGATTCTTTTGTTCATCA 171

Db 311 LeuAspSerPheValPheSer 318

RESULT 15
T27257

hypothetical protein Y62H9A.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T27257
R;Steward, C.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20334
A;Accession: T27257
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-366 <WIL>
A;Cross-references: EMBL:AL032630; PIDN:CAA21558.1; GSPDB:GN00028; CESP:Y62H9A.1
A;Experimental source: clone Y62H9A
C;Genetics:
A;Gene: CESP:Y62H9A.1
A;Map position: X
A;Introns: 25/3; 206/3; 287/2; 342/1

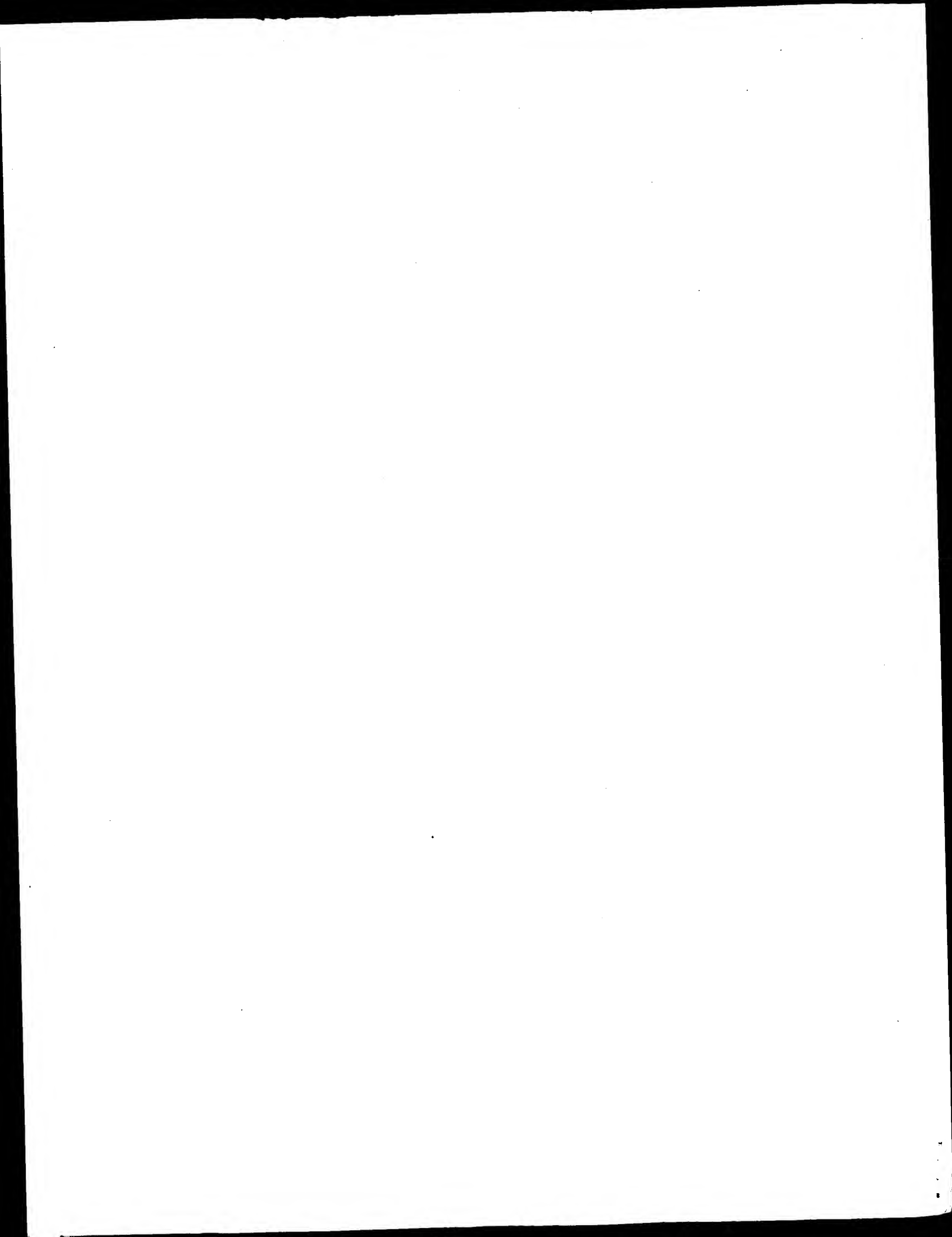
Alignment Scores:
Pred. No.: 44.4 Length: 366
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x T27257 (1-366)

QY 2082 TTTTCTGTATATGGGTTTGTGA 2105

Db 71 PheSerValTyrMetGlyPheVal 78

Search completed: April 16, 2003, 16:41:50
Job time : 48.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:26:10 ; Search time 25.5 Seconds
(without alignments)
7283.569 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 732
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_091001_21192/app_query.fasta_1.2375
-DB=SwissProt_40 -QFWT=fastan -SUFFIX=olin2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_1_24@runat_04042003_091001_21192 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	466	63.7	501	1 XCT_HUMAN	Q9upy5 homo sapien
2	46	6.3	502	1 XCT_MOUSE	Q9wtr6 mus musculus
3	12	1.6	511	1 YLAI_HUMAN	Q9um01 homo sapien
4	8	1.1	67	1 RS21_AQUAE	O67028 aquifex ae
5	8	1.1	82	1 COAB_BPPF1	P03621 bacterioph
6	8	1.1	260	1 FAPR_ECOLI	P23774 escherichia
7	8	1.1	354	1 AC14_HUMAN	P35250 homo sapien
8	8	1.1	359	1 AC14_CHICK	P53033 gallus gall
9	8	1.1	419	1 GSC_DROME	P54366 drosophila
10	8	1.1	467	1 2A5E_HUMAN	Q16537 h serine/th
11	8	1.1	499	1 AMPA_CHLPN	Q9z8f8 chlamydia p
12	8	1.1	507	1 LATI_HUMAN	Q01650 homo sapien
13	8	1.1	512	1 LATI_MOUSE	Q9z127 mus musculus
14	8	1.1	512	1 LATI_RAT	Q63016 rattus norv
15	8	1.1	531	1 LAT2_MOUSE	Q9qwx9 mus musculus
16	8	1.1	533	1 LAT2_RAT	Q9wvr6 rattus norv
17	8	1.1	535	1 LAT2_HUMAN	Q9uh15 homo sapien
18	8	1.1	546	1 TCPE_SCHPO	Q9utm4 schizosacch

19	8	1.1	554	1	REC_N_RICPR	Q9zdy2 rickettsia
20	8	1.1	786	1	ST5B_MOUSE	P42232 mus musculus
21	8	1.1	786	1	ST5B_RAT	P52632 rattus norv
22	8	1.1	793	1	ST5A_MOUSE	P42230 mus musculus
23	8	1.1	793	1	ST5A_RAT	Q62771 rattus norv
24	8	1.1	1447	1	SGS1_YEAST	P35187 saccharomyc
25	8	1.1	1483	1	UFD4_YEAST	P33202 saccharomyc
26	8	1.1	2226	1	POLG_HPAV2	P26580 hepatitis a
27	8	1.1	2226	1	POLG_HPAV4	P26581 hepatitis a
28	8	1.1	2226	1	POLG_HPAV8	P26582 hepatitis a
29	8	1.1	2227	1	POLG_HPAVH	P08617 hepatitis a
30	8	1.1	2227	1	POLG_HPAVL	P06441 hepatitis a
31	7	1.0	45	1	Y843_HAEIN	P44059 haemophilus
32	7	1.0	56	1	Y048_BORBU	O51077 borrelia bu
33	7	1.0	104	1	C552_BRAJA	P24059 bradyrhizob
34	7	1.0	108	1	CYC_CURLU	Q96vp3 curvularia
35	7	1.0	112	1	CYC_ARATH	P29380 arabidopsis
36	7	1.0	127	1	SECE_ECOLI	P16920 escherichia
37	7	1.0	127	1	SECE_SALTY	Q919k1 salmonella
38	7	1.0	129	1	RS9_CHLTR	O84128 chlamydia t
39	7	1.0	130	1	INI2_HUMAN	P09912 homo sapien
40	7	1.0	130	1	INI2_PANTR	Q28808 pan troglod
41	7	1.0	133	1	RS9_CHLMU	Q9pkx2 chlamydia m
42	7	1.0	133	1	SECR_MOUSE	Q08535 mus musculus
43	7	1.0	134	1	RS9_CHLPN	Q9z8t8 chlamydia p
44	7	1.0	135	1	ATPE_SYNY3	P26533 synechocyst
45	7	1.0	141	1	X_WHV1	P03167 woodchuck h

ALIGNMENTS

RESULT 1					
XCT_HUMAN					
ID	XCT_HUMAN	STANDARD;	PRT;	501 AA.	
AC	Q9UPY5;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Cystine/glutamate transporter (Amino acid transport system xc-) (xCT)				
DE	(Calcium channel blocker resistance protein CCBRL)				
GN	SLC7A11				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21082210; PubMed=11213471;				
RA	Sato H., Tamba M., Kuriyama-Matsumura K., Okuno S., Bannai S.;				
RT	"Molecular cloning and expression of human xCT, the light chain of				
RT	amino acid transport system xc-";				
RL	Antioxid. Redox. Signal. 2:665-671(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Conklin D.S., Beach D.H.;				
RT	"CCBR1, novel CD98 light chain implicated in redox control and calcium				
RT	signaling.";				
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta;				
RA	Chancy C.D., Kekuda R., Wang H., Huang W., Prasad P.D., Smith S.B.,				
RA	Ganapathy V.;				
RT	"Structure, function and regulation of human cystine/glutamate				
RT	transporter in retinal pigment epithelial cells.";				
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Borsani G., Manzoni M., Palacin M., Pineda M., Gasol E.;				
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				

RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY EXCHANGE OF ANIONIC
CC AMINO ACIDS WITH HIGH SPECIFICITY FOR ANIONIC FORM OF CYSTINE AND
CC GLUTAMATE (BY SIMILARITY).
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLC3A2/4F2HC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB026891; BAA82628.1; -.
CC EMBL; AF200708; AAG35592.1; -.
CC EMBL; AF252872; AAK49111.1; -.
CC EMBL; AJ277882; CAC81905.1; -.
CC EMBL; BC012087; AAH12087.1; -.
CC Genew; HGNC:11059; SLC7A11.
CC InterPro; IPR002293; AA/rel_pmease1.
CC InterPro; IPR004760; L_AA_transport.
CC InterPro; IPR004841; Permease.
CC Pfam; PF00324; aa_permeases; 1.
CC TIGRFAMs; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 423 443 POTENTIAL.
FT TRANSMEM 450 470 POTENTIAL.
FT CARBOHYD 314 314 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 501 AA; 55422 MW; 3EF2648B94A9F59E CRC64;

US-09-667-170A-440 (1-2239) x XCT_HUMAN (1-501)

Alignment Scores:
Pred. No.: 0 Length: 501
Score: 466.00 Matches: 466
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 63.66% Indels: 0
DB: 1 Gaps: 0

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTAGGGGAGTCTCCATTATCATTTGGC 393
Db 36 GluLysValGlnLeuLysArgLysValThrLeuLeuArgGlyValSerIleIleGly 55
QY 394 ACCATCATTTGAGCAGGAATCTTCATCTCTCTAAGGGGTGCTCCAGAACACGGGCAGC 453
Db 56 ThrIleIleGlyAlaGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer 75
QY 454 GTGGGCATGCTCTGACCATCTGGAGGGTGTGTGGGGTCTGTCACTATTGGAGCTTTG 513
Db 76 ValGlyMetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeu 95
QY 514 TCTTATGCTGAATGGGAACAACATAPAAAGAAATCTGGAGGTCATTACACATATTTTG 573
Db 96 SerTyrAlaGluLeuGlyThrThrIleLysLysSerGlyGlyHisTyrThrTyrIleLeu 115

QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGTGGAACTCCTCATAATACGC 633
Db 116 GluValPheGlyProLeuProAlaPheValArgValTrpValGluLeuIleArg 135
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACCTACATTCTGGAACCATTTT 693
Db 136 ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProphePhe 155
QY 694 ATTCAATGTCAAAATCCCTGAACCTGCGATCAAGCTCATACAGCTGTGGGCATACTGTA 753
Db 156 IleGlnCysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrVal 175
QY 754 GTGATGGTCTTAAATAGCATGAGTGTACAGCTGGAGCGCCGGATCCAGATTTTCTTAACC 813
Db 176 ValMetValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThr 195
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAAATATAGTCCCTGGAGTTATGCAGCTAATATAA 873
Db 196 PheCysLysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLys 215
QY 874 GGTCAAAACGCAGAACTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAGATTTACGGCGTTG 933
Db 216 GlyGlnThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerIleThrArgLeu 235
QY 934 CCACTGGCTTTTATATGGAATGTATGCATATGCTGGCTGGTGTTCCTCAACATTTGTT 993
Db 236 ProLeuAlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheVal 255
QY 994 ACTGAAGAAGTAGAAACCCCTGAAAAACCATTCCTTGGCAATATGATATATCCATGGCC 1053
Db 256 ThrGluGluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAla 275
QY 1054 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCTTACTTTACGACCATTAATGCTGAG 1113
Db 276 IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrIleAsnAlaGlu 295
QY 1114 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 296 GluLeuLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe 315
QY 1174 TCATTAGCAGTTCGGATCTTTGTTGGCTCTCTCTCTGCTTGGCTCCATGAACGGTGTGTG 1233
Db 316 SerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGlyVal 335
QY 1234 TTTGCTGTCTCCAGTTATTCTATGTTGGCTCTCTCTCTGAGGGTCCACCTTCCAGAAATCCTC 1293
Db 336 PheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeu 355
QY 1294 TCCATGATTCATGTCGCAAGCACACTCTCTACAGCTGTTATTGTTTGGCACCCCTTTG 1353
Db 356 SerMetIleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeu 375
QY 1354 ACAATGATAATGCTCTCTCTGAGACCTCGACAGTCTTTTGAATTTCTCCTCAGTTTGGC 1413
Db 376 ThrMetIleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAla 395
QY 1414 AGGTGGCTTTTATTTGGGCTGGCAGTGTGCTGGGCTGATTATCTTCGATACAAATGCCCA 1473
Db 396 ArgTrpLeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysPro 415
QY 1474 GATATGCATCGTCTCTCAAGGTGCCACTGTTTATCCAGCTTGTTCCTTCATCATGC 1533
Db 416 AspMetHisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCys 435
QY 1534 CTCTTCATGGTGGCTCTCTCTCTATTCGGACCCCATTTAGTACAGGGATTGGCTTCGTC 1593
Db 436 LeuPheMetValAlaLeuSerLeuTyrSerAspProPheSerThrGlyIleGlyPheVal 455
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGGACAAGAAACCCAGG 1653
Db 456 IleThrLeuThrGlyValProAlaTyrTyrLeuPheIleIleThrAspLysLysProArg 475
QY 1654 TGGTTTAGAATAATGTACAGAGAAAATAACCCAGAACATTAACAATAATACTGGAAGTTGTA 1713


```
Db 476 TrpPheArgIleMetSerGluLysIleThrArgThrLeuGlnIleLeuGluValVal 495
|||||
QY 1714 CCAGACAGATAAGTTA 1731
|||||
Db 496 ProGluGluAspLysLeu 501
|||||
RESULT 2
XCT_MOUSE
ID XCT_MOUSE STANDARD; PRT; 502 AA.
AC Q9WTR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cystine/glutamate transporter (Amino acid transport system xc-) (xCT).
GN SLC7A11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99223452; PubMed=10206947;
RA Sato H., Tamba M., Ishii T., Bannai S.;
RT "Cloning and expression of a plasma membrane cystine/glutamate
exchange transporter composed of two distinct proteins.";
RL J. Biol. Chem. 274:11455-11458(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ; TISSUE=Liver;
RA Sasaki H., Sato H., Bannai S.;
RT "Isolation and functional characterization of mouse cystine/glutamate
exchange transporter gene.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY EXCHANGE OF ANIONIC
CC AMINO ACIDS WITH HIGH SPECIFICITY FOR ANIONIC FORM OF CYSTINE AND
CC GLUTAMATE.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB022345; BAA77220.1; -
DR EMBL; AB037661; BAA90522.1; -
DR EMBL; AB037650; BAA90522.1; JOINED.
DR EMBL; AB037651; BAA90522.1; JOINED.
DR EMBL; AB037652; BAA90522.1; JOINED.
DR EMBL; AB037653; BAA90522.1; JOINED.
DR EMBL; AB037654; BAA90522.1; JOINED.
DR EMBL; AB037655; BAA90522.1; JOINED.
DR EMBL; AB037656; BAA90522.1; JOINED.
DR EMBL; AB037657; BAA90522.1; JOINED.
DR EMBL; AB037658; BAA90522.1; JOINED.
DR EMBL; AB037659; BAA90522.1; JOINED.
DR EMBL; AB037660; BAA90522.1; JOINED.
DR MGD; MGI:1347355; SLC7a11.
DR InterPro; IPR002293; AA/rele_prmeasel.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 44 64 POTENTIAL.
```

```
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 423 443 POTENTIAL.
FT TRANSMEM 450 470 POTENTIAL.
FT CARBOHYD 19 19 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 502 AA; 55455 MW; A4A185102D83992A CRC64;

Alignment Scores:
Pred. No.: 2.02e-39 Length: 502
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.28% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x XCT_MOUSE (1-502)

QY 1171 TTCTCATTAGCAGTCCGATCTTTGTTGGCCTCTCCTGCTTGGTCCATGAACGGTGGT 1230
|||||
Db 315 PheSerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGly 334

QY 1231 GTGTTTGCTCTCCAGGTTATTCTATGTTCGCTCTCGAGAGGGTCACCTTCCAGAAATC 1290
|||||
Db 335 ValPheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIle 354

QY 1291 CTCTCCATGATTCATGTC 1308
|||||
Db 355 LeuSerMetIleHisVal 360

RESULT 3
YLAL_HUMAN
ID YLAL_HUMAN STANDARD; PRT; 511 AA.
AC Q9UM01; O9P2V5; O95984;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Y+L amino acid transporter 1 (Y+L-type amino acid transporter 1)
DE (Y+LAT-1) (Y+LAT1) (Monocyte amino acid permease 2) (MOP-2).
GN SLC7A7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.; AND VARIANT LPI ARG-334.
RX MEDLINE=99047611; PubMed=9829974;
RA Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J.,
RA Shi Y.-B., Zorzano A., Palacin M.;
RT "Identification and characterization of a membrane protein (Y+L amino
acid transporter-1) that associates with 4F2hc to encode the amino
acid transport activity Y+L. A candidate gene for lysinuric protein
intolerance.";
RN J. Biol. Chem. 273:32437-32445(1998).
RP [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=99094891; PubMed=9878049;
RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
RT "Amino acid transport of Y+L-type by heterodimers of 4F2hc/CD98 and
members of the glycoprotein-associated amino acid transporter
family.";
RL EMBO J. 18:49-57(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
```

RX MEDLINE=99178268; PubMed=10080183;
RA Borsani G., Bassi M.T., Sperandio M.P., De Grandi A., Buoninconti A.,
RA Riboni M., Manzoni M., Incerti B., Pepe A., Andria G., Ballabio A.,
RA Sebastio G.;
RT "SLC7A7, encoding a putative permease-related protein, is mutated in
RT patients with lysinuric protein intolerance.";
RL Nat. Genet. 21:297-301(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Fukasawa Y., Segawa H., Endou H., Kanai Y.;
RT "Characterization of a human system y+L amino acid transporters.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20202143; PubMed=10737982;
RA Noguchi A., Shoji Y., Koizumi A., Takahashi T., Shoji Y.,
RA Matsumori M., Kayo T., Ohata T., Wada Y., Yoshimura I., Maisawa S.,
RA Konishi M., Takasago Y., Takada G.;
RT "SLC7A7 genomic structure and novel variants in three Japanese
RT lysinuric protein intolerance families.";
RL Hum. Mutat. 15:367-372(2000).
RN [6]
RP SEQUENCE FROM N.A.
RA Takayama K., Yoshimoto M.;
RT "Molecular and biological characterization of a novel monocyte amino
RT acid permease, MOP-2.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RP VARIANT LPI ARG-386.
RX MEDLINE=20100611; PubMed=106311139;
RA Sperandio M.P., Bassi M.T., Riboni M., Parenti G., Buoninconti A.,
RA Manzoni M., Incerti B., Larocca M.R., Di Rocco M., Strisciuglio P.,
RA Dianzani I., Parini R., Candido M., Endo F., Ballabio A., Andria G.,
RA Sebastio G., Borsani G.;
RT "Structure of the SLC7A7 gene and mutational analysis of patients
RT affected by lysinuric protein intolerance.";
RL Am. J. Hum. Genet. 66:92-99(2000).
RN [9]
RP VARIANTS LPI VAL-54 AND ASP-338.
RX MEDLINE=20122253; PubMed=10655553;
RA Mykkaenen J., Torrents D., Pineda M., Camps M., Yoldi M.E.,
RA Horelli-Kuitunen N., Huoponen K., Heinonen M., Oksanen J., Simell O.,
RA Savontaus M.-L., Zorzano A., Palacin M., Aula P.;
RT "Functional analysis of novel mutations in y+LAT-1 amino acid
RT transporter gene causing lysinuric protein intolerance (LPI).";
RL Hum. Mol. Genet. 9:431-438(2000).
CC -!- FUNCTION: SODIUM-INDEPENDENT EXCHANGER OF CATIONIC AND LARGE
CC NEUTRAL AMINO ACIDS.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY >> PERIPHERAL BLOOD
CC LEUKOCYTES >> LUNG > PLACENTA = SPLEEN > SMALL INTESTINE.
CC -!- DISEASE: DEFECTS IN SLC7A7 ARE A CAUSE OF LYSINURIC PROTEIN
CC INTOLERANCE (LPI). LPI IS AN AUTOSOMAL RECESSIVE MULTISYSTEM
CC DISORDER FOUND MAINLY IN FINLAND AND ITALY. ON A NORMAL DIET, LPI
CC PATIENTS PRESENT POOR FEEDING, VOMITING, DIARRHEA, EPISODES OF
CC HYPERAMMONAEMIC COMA AND GROWTH RETARDATION. HEPATOSPLENOMEGALY,
CC OSTEOPOROSIS AND A LIFE-THREATENING PULMONARY INVOLVEMENT
CC (ALVEOLAR PROTEINOSIS) ARE ALSO SEEN. BIOCHEMICALLY LPI IS
CC CHARACTERIZED BY A DEFECT IN THE PLASMA MEMBRANE TRANSPORT OF
CC DIBASIC AMINO ACIDS.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y18474; CAB40136.1; -.
DR EMBL; AJ130718; CAA10198.1; -.
DR EMBL; AF092032; AAC83706.1; -.
DR EMBL; AB020532; BAA87623.1; -.
DR EMBL; AB031537; BAA95120.1; ALT_SEQ.
DR EMBL; AB031529; BAA95120.1; JOINED.
DR EMBL; AB031530; BAA95120.1; JOINED.
DR EMBL; AB031531; BAA95120.1; JOINED.
DR EMBL; AB031532; BAA95120.1; JOINED.
DR EMBL; AB031533; BAA95120.1; JOINED.
DR EMBL; AB031534; BAA95120.1; JOINED.
DR EMBL; AB031535; BAA95120.1; JOINED.
DR EMBL; AB031536; BAA95120.1; JOINED.
DR EMBL; AB011263; BAB11849.1; -.
DR EMBL; BC003062; AAH03062.1; -.
DR EMBL; BC010107; AAH10107.1; -.
DR Genew; HGNC:11065; SLC7A7.
DR MIM; 603593; -.
DR MIM; 222700; -.
DR InterPro; IPR002293; AA/rel_primease1.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
KW Transport; Amino-acid transport; Transmembrane; Disease mutation.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 222 242 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 383 403 POTENTIAL.
FT TRANSMEM 416 436 POTENTIAL.
FT TRANSMEM 441 461 POTENTIAL.
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 54 54 G -> V (IN LPI).
FT VARIANT 334 334 /FTid=VAR_010261.
FT VARIANT 338 338 L -> R (IN LPI).
FT VARIANT 386 386 /FTid=VAR_010262.
FT VARIANT 386 386 G -> D (IN LPI).
FT CONFLICT 91 91 /FTid=VAR_010999.
FT CONFLICT 91 91 S -> R (IN LPI).
FT CONFLICT 91 91 /FTid=VAR_011000.
FT CONFLICT 91 91 A -> V (IN REF. 4).
SQ SEQUENCE 511 AA; 55990 MW; A71D677B6B075894 CRC64;
Alignment Scores:
Pred. No.: 0.00109 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 1 Gaps: 0
US-09-667-170A-440 (1-2239) x YLAI_HUMAN (1-511)
QY 517 TATGCTGAATGGGAACAACACTATAAGAAATCTGGA 552
|||||
Db 90 TyrAlaGluLeuGlyThrThrIleLysSergly 101
RESULT 4
RS21_AQUAE
ID RS21_AQUAE STANDARD; PRT; 67 AA.
AC O67028;
DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN 30S ribosomal protein S21.
OS RPSU OR AQ_867 OR AQ_867A.
OC Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000711; AAC06990.1; -;
DR InterPro; IPR001911; Ribosomal_S21.
DR Pfam; PF01165; Ribosomal_S21; 1.
DR PRINTS; PR00976; RIBOSOMALS21.
DR ProDom; PD005521; Ribosomal_S21; 1.
DR TIGRFAMs; TIGR00030; S21p; 1.
DR PROSITE; PS01181; RIBOSOMAL_S21; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 67 AA; 8261 MW; E1897087A487EF70 CRC64;

Alignment Scores:
Pred. No.: 24 Length: 67
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x RS21_AQUAE (1-67)
QY 77 GAAAGAAAGAAAAGAGAGAGAGG 100
|||||
Db 43 GluArgLysLysLysLysLysArgGluArg 50

RESULT 5
COAB_BPPF1
ID COAB_BPPF1 STANDARD; PRT; 82 AA.
AC P03621;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein B precursor (Major coat protein).
GN VIII.
OS Bacteriophage Pfl.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=10871;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25102-B1;
RX MEDLINE=91186399; PubMed=2010913;
RA Hill D.F., Short N.J., Perham R.N., Petersen G.B.;
RT "DNA sequence of the filamentous bacteriophage Pfl.";
RL J. Mol. Biol. 218:349-364(1991).
RN [2]
RP SEQUENCE OF 37-82.

RX MEDLINE=75082355; PubMed=1110754;
RA Nakashima Y., Wiseman R.L., Konigsberg W., Marvin D.A.;
RT "Primary structure and sidechain interactions of pfl filamentous
RT bacterial virus coat protein.";
RL Nature 253:68-71(1975).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=94310463; PubMed=8036516;
RA Liu D.J., Day L.A.;
RT "Pfl virus structure: helical coat protein and DNA with paraxial
RT phosphates.";
RL Science 265:671-674(1994).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=92022521; PubMed=1925543;
RA Nambudripad R., Stark W., Opella S.J., Makowski L.;
RT "Membrane-mediated assembly of filamentous bacteriophage Pfl coat
RT protein.";
RL Science 252:1305-1308(1991).
CC -!- FUNCTION: COAT PROTEIN B IS THE MAJOR COAT PROTEIN OF THE VIRION.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52107; CAA36331.1; -;
DR PIR; A04229; VCBPPF.
DR PIR; S15143; S15143.
DR PIR; S20699; S20699.
DR PDB; 1IFM; 31-JUL-94.
DR PDB; 1IFN; 31-JUL-94.
DR PDB; 2IFN; 01-JAN-96.
DR PDB; 2IFN; 01-JAN-96.
DR PDB; 3IFM; 01-JAN-96.
DR PDB; 4IFM; 01-JAN-96.
DR PDB; 1PFI; 26-JAN-95.
KW Coat protein; Signal; 3D-structure; Transmembrane.
FT SIGNAL 1 36
FT CHAIN 37 82 COAT PROTEIN B.
FT DOMAIN 37 55 EXTRACELLULAR.
FT TRANSMEM 56 78
FT DOMAIN 79 82 INTRACELLULAR.
FT CONFLICT 45 45 E -> Q (IN REF. 2).
SQ SEQUENCE 82 AA; 8377 MW; 25FD055DA4661504 CRC64;

Alignment Scores:
Pred. No.: 23.2 Length: 82
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x COAB_BPPF1 (1-82)
QY 1432 CTGGCAGTTGCTGGGCTGATTAT 1455
|||||
Db 69 LeuAlaValAlaGlyLeuIleTyr 76

RESULT 6
FAPR_ECOLI
ID FAPR_ECOLI STANDARD; PRT; 260 AA.
AC P23774;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 987P fimbrial operon positive regulatory protein fapR.
GN FAPR.

OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91171879; PubMed=2077360;
 RA Klaasen P., de Graaf F.K.;
 RT "Characterization of FapR, a positive regulator of expression of the
 987P operon in enterotoxigenic Escherichia coli.";
 RL Mol. Microbiol. 4:1779-1783(1990).
 CC -|- FUNCTION: POSITIVE REGULATOR OF THE EXPRESSION OF THE 987P OPERON
 CC FOR THE FIMBRIAL PROTEIN IN ENTEROTOXIGENIC ESCHERICHIA COLI.
 CC -|- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X53494; CAA37578.1; -.
 DR PIR; S11984; S11984.
 DR InterPro; IPR000005; HTHARAC.
 DR Pfam; PF00165; HTH_Arac; 2.
 DR PRINTS; PR00032; HTHARAC.
 DR SMART; SM00342; HTH_ARAC; 1.
 DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
 DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
 DR Transcription regulation; DNA-binding; Activator.
 KW DNA_BIND 170 189 H-T-H MOTIF (BY SIMILARITY).
 FT SEQUENCE 260 AA; 30349 MW; 1F73401737062C7A CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 19.4 Length: 260
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0
 US-09-667-170A-440 (1-2239) x FAPR_ECOLI (1-260)
 QY 1135 GTGGCAGTGACCTTTCTGAGCGG 1158
 Db 148 ValAlaValThrPheSerGluArg 155
 RESULT 7
 AC14_HUMAN
 ID AC14_HUMAN STANDARD; PRT; 354 AA.
 AC P35250; P32846;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Activator 1 40 kDa subunit (Replication factor C 40 kDa subunit) (A1
 DE 40 kDa subunit) (RF-C 40 kDa subunit) (RFC40).
 GN RFC2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 232-249.
 RX MEDLINE=92212860; PubMed=1313560;
 RA Chen M., Pan Z.-Q., Hurwitz J.;
 RT "Sequence and expression in Escherichia coli of the 40-kDa subunit of
 activator 1 (replication factor C) of HeLa cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2516-2520(1992).
 RN [2]
 RP REVISIONS.

RA Hurwitz J.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Martindale D.W., Koop B.F.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE
 CC DELTA AND EPSILON REQUIRES THE ACTION OF THE ACCESSORY PROTEINS
 CC PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) AND ACTIVATOR 1. THE
 CC 40 kDa SUBUNIT BINDS ATP.
 CC -|- SUBUNIT: HETEROPENTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND
 CC 36.5 kDa THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP.
 CC -|- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -|- DISEASE: Haploinsufficiency of RFC2 may be the cause of certain
 CC cardiovascular and musculo-skeletal abnormalities observed in
 CC Williams-Beuren syndrome (WBS), a rare developmental disorder. It
 CC is a contiguous gene deletion syndrome involving genes from
 CC chromosome band 7q11.23.
 CC -|- SIMILARITY: BELONGS TO THE ACTIVATOR 1 36 TO 40 kDa SUBUNITS
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M87338; AAB09786.1; -.
 DR EMBL; AF045555; AAC04860.1; -.
 DR PIR; A42700; A42700.
 DR Genew; HGNC:9970; RFC2.
 DR MIM; 600404; -.
 DR MIM; 194050; -.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_cent.
 DR InterPro; IPR000862; RFCdomain.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 KW DNA replication; ATP-binding; Nuclear protein;
 KW Williams-Beuren syndrome.
 FT NP_BIND 76 83 ATP (POTENTIAL).
 FT CONFLICT 244 244 G -> L (IN REF. 1; AA SEQUENCE).
 SQ SEQUENCE 354 AA; 39157 MW; B50AC8EEF89F64A9 CRC64;
 Alignment Scores:
 Pred. No.: 18.5 Length: 354
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0
 US-09-667-170A-440 (1-2239) x AC14_HUMAN (1-354)
 QY 1761 CTGCCCCAAGGGGAGACACAAATA 1784
 Db 130 LeuProLysGlyArgHisLysile 137
 RESULT 8
 AC14_CHICK
 ID AC14_CHICK STANDARD; PRT; 359 AA.
 AC P53033;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Activator 1 40 kDa subunit (Replication factor C 40 kDa subunit) (A1
 DE 40 kDa subunit) (RF-C 40 kDa subunit) (RFC40).
 GN RFC2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;


```
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Leghorn; TISSUE=Cochlea;
RX MEDLINE=94347215; PubMed=8068208;
RA Oberholtzer J., Cohen E.L., Davis J.G.;
RT "Molecular cloning of a chick cochlea cDNA encoding a subunit of DNA
RT replication factor C/activator 1.";
RL DNA Cell Biol. 13:857-863(1994).
CC -!- FUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE
CC DELTA AND EPSILON REQUIRES THE ACTION OF THE ACCESSORY PROTEINS
CC PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) AND ACTIVATOR 1. THE
CC 40 kDa SUBUNIT BINDS ATP (BY SIMILARITY).
CC -!- SUBUNIT: HETEROPENTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND
CC 36.5 kDa THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE ACTIVATOR 1 36 TO 40 kDa SUBUNITS
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U12438; AAA20552.1; -.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR003959; AAA_ATPase_cent.
CC DR InterPro; IPR000862; RFCdomain.
CC DR Pfam; PF00004; AAA; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR DNA replication; ATP-binding; Nuclear protein.
KW NP_BIND 81 88 ATP (POTENTIAL).
FT SEQUENCE 359 AA; 39706 MW; A2FD38F964CD11F9 CRC64;
SQ -----
Alignment Scores:
Pred. No.: 18.4 Length: 359
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0
US-09-667-170A-440 (1-2239) x AC14_CHICK (1-359)
QY 1761 CTGCCCCAAGGGGAGACACAAAATA 1784
Db 135 LeuProLysGlyArgHisLysile 142
RESULT 9
GSC_DROME
ID GSC_DROME STANDARD; PRT; 419 AA.
AC P54366; Q9VPR9;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein goosecoid.
GN GSC OR CG2851.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=96202483; PubMed=8625850;
RA Goriely A., Stella M., Coffinier C., Kessler D., Mailhos C.,
RA Dessain S., Desplan C.;
RT "A functional homologue of goosecoid in Drosophila.";
RL Development 122:1641-1650(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96272167; PubMed=8670808;
RA Hahn M., Jackle H.;
RT "Drosophila goosecoid participates in neural development but not in
RT body axis formation.";
RL EMBO J. 15:3077-3084(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nusskern D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: APPEARS TO REGULATE REGIONAL DEVELOPMENT OF SPECIFIC
CC TISSUES. CAN RESCUE AXIS POLARITY IN UV-RADIATED XENOPUS
CC EMBRYOS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: IN EARLY EMBRYO DEVELOPMENT, EXPRESSION
CC CONFINED TO TWO REGIONS; A HORSESHOE-LIKE PATTERN ACROSS THE
CC DORSAL SIDE WHICH IS DESTINED TO FORM THE BRAIN HEMISPHERES AND A
CC SECOND DOMAIN WHICH INVAGINATES INSIDE THE STOMODEUM AND WHICH, IS
CC FATED TO FORM THE FOREGUT, RING GLAND AND STOMATOGASTRIC NERVOUS
CC SYSTEM (SNS).
CC -!- SIMILARITY: BELONGS TO THE PAIRED HOMEBOX FAMILY.
CC "BICOID" SUBFAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

DR EMBL; X95420; CAA64699.1; -
DR EMBL; U52968; AAB17948.1; -
DR EMBL; AE003589; AAF51473.1; -
DR HSSP; P06601; 1FJL.
DR TRANSFAC; T04041; -
DR FlyBase; FBgn0010323; Gsc.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Developmental protein; Nuclear protein; DNA-binding; Homeobox.
FT DOMAIN 104 107 POLY-ALA.
FT DOMAIN 164 169 POLY-SER.
FT DOMAIN 195 199 POLY-ALA.
FT DNA_BIND 286 345 HOMEBOX.
SQ SEQUENCE 419 AA; 44949 MW; 851A4C46AA861FB9 CRC64;

Alignment Scores:
Pred. No.: 18 Length: 419
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x GSC_DROME (1-419)

QY 1724 TCTTCTTCTGTACAACTTCCAGT 1701

Db 370 SerSerSerGlyThrThrSerSer 377

RESULT 10

2A5E_HUMAN

ID 2A5E_HUMAN STANDARD; PRT; 467 AA.
AC Q16537;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit,
DE epsilon isoform (PP2A, B subunit, B' epsilon isoform) (PP2A, B
DE subunit, B56 epsilon isoform) (PP2A, B subunit, PP61 epsilon isoform)
DE (PP2A, B subunit, R5 epsilon isoform).
GN PPP2R5E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 449-455.
RC TISSUE=Fetal retina;
RX MEDLINE=96276417; PubMed=8694763;
RA Zolnierowicz S., van Hoof C., Andjelkovic N., Cron P., Stevens I.,
RA Merlevede W., Goris J., Hemmings B.A.;
RT "The variable subunit associated with protein phosphatase 2A0 defines
a novel multimer family of regulatory subunits.";
RL Biochem. J. 317:187-194(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96355607; PubMed=8703017;
RA McCright B., Rivers A.M., Audlin S., Virshup D.M.;
RT "The B56 family of protein phosphatase 2A (PP2A) regulatory subunits
encodes differentiation-induced phosphoproteins that target PP2A to
both nucleus and cytoplasm.";
RL J. Biol. Chem. 271:22081-22089(1996).
CC -!- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
CC COMPARTMENT.
CC -!- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa

CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
CC AND CELL SIGNALING MOLECULES.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: PHOSPHORYLATED ON SERINE RESIDUES.
CC -!- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z69029; CAA93153.1; -
DR EMBL; L76703; AAB69752.1; -
DR Genew; HGNC:9313; PPP2R5E.
DR MIM; 601647; -
DR InterPro; IPR002554; B56.
DR Pfam; PF01603; B56; 1.
KW Phosphorylation; Multigene family.
SQ SEQUENCE 467 AA; 54699 MW; DD9CE11433F499CF CRC64;

Alignment Scores:

Pred. No.: 17.7 Length: 467
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x 2A5E_HUMAN (1-467)

QY 79 AAGAAAGAAAAAGAGAGAGAGAGAA 102

Db 438 LysLysGluLysGluArgGluGlu 445

RESULT 11

AMPA_CHLPN

ID AMPA_CHLPN STANDARD; PRT; 499 AA.
AC Q928F8; Q9JQJ2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE (LAP) (Leucyl aminopeptidase).
GN PEPA OR CPN0385 OR CP0370.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia

RT pneumonias AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CWL029 from USA."
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -!- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR
 CC TURNOVER OF INTRACELLULAR PROTEINS. CATALYZES THE REMOVAL OF
 CC UNSUBSTITUTED N-TERMINAL AMINO ACIDS FROM VARIOUS PEPTIDES (BY
 CC SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-1-
 CC Xbb-, in which Xaa is preferably Leu, but may be other amino acids
 CC including Pro although not Arg or Lys, and Xbb may be Pro.
 CC -!- COFACTOR: MANGANESE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AE001623; AAD18529.1; -;
 DR EMBL; AE002199; AAF38219.1; -;
 DR EMBL; AP002546; BAA98593.1; -;
 DR HSP; P00727; 1BPN.
 DR MEROPS; M17; UPW; -;
 DR PHCI-2DPAGE; Q9Z8F8; -;
 DR TIGR; CP0370; -;
 DR InterPro; IPR000819; Peptidase_M17.
 DR Pfam; PF00883; Peptidase_M17; 1.
 DR Pfam; PF02789; Peptidase_M17_N; 1.
 DR PRINTS; PR00481; LAMNOPPTDASE.
 DR PROSITE; PS00631; CYTOSOLAP; 1.
 KW Hydrolase; Amino-peptidase; Manganese; Complete proteome.
 FT METAL 263 MANGANESE 2 (BY SIMILARITY).
 FT METAL 268 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT METAL 286 MANGANESE 2 (BY SIMILARITY).
 FT METAL 345 MANGANESE 1 (BY SIMILARITY).
 FT METAL 347 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT ACT_SITE 275 POTENTIAL.
 FT ACT_SITE 349 POTENTIAL.
 SQ SEQUENCE 499 AA; 54509 MW; B0F281B1DF4B9EC2 CRC64;

Alignment Scores:
 Pred. No.: 17.5 Length: 499
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x AMPA_CHLPN (1-499)

QY 380 GAGACTCCCTCAGTAAAGTGACT 357
 |||||
 Db 149 GluThrProLeuSerLysValThr 156

RESULT 12
 LATI_HUMAN
 ID LATI_HUMAN STANDARD; PRT; 507 AA.
 AC Q01650; Q9UBN8; Q9UP15; Q9UQC0;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Large neutral amino acids transporter small subunit 1 (L-type amino
 DE acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC) (CD98 light
 DE chain) (Integral membrane protein El6) (hLAT1).
 GN SLC7A5 OR LAT1 OR MPE16 OR CD98LC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98421678; PubMed=9751058;
 RA Mastrobardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffing J.,
 RA Shoemaker C.B., Verrey F.;
 RT "Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a
 RT permease family.";
 RL Nature 395:288-291(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Yanagida O., Segawa H., Miyamoto K., Takeda E., Goya T., Endou H.,
 RA Kanai Y.;
 RT "Cloning and characterization of a human system L amino acid
 RT transporter.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99160855; PubMed=10049700;
 RA Prasad P.D., Wang H., Huang W., Kekuda R., Rajan D.P., Leibach F.H.,
 RA Ganapathy V.;
 RT "Human LAT1, a subunit of system L amino acid transporter: molecular
 RT cloning and transport function.";
 RL Biochem. Biophys. Res. Commun. 255:283-288(1999).
 RN [4]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=99172172; PubMed=10072483;
 RA Tsudome M., Ito M., Takebayashi S., Okumura K., Nishio M.,
 RA Kawano M., Kusagawa S., Komada H., Ito Y.;
 RT "Primary structure of the light chain of fusion regulatory protein-
 RT 1/CD98/4F2 predicts a protein with multiple transmembrane domains that
 RT is almost identical to the amino acid transporter El6.";
 RL J. Immunol. 162:2462-2466(1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Minato N., Iwai K., Takizawa C., Nakamura E.;
 RT "Human 4F2 light chain: amino acid transporter.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 267-507 FROM N.A.
 RC TISSUE=Peripheral blood lymphocytes;
 RX MEDLINE=92283834; PubMed=1597461;
 RA Gaugitsch H.W., Prieschl E.E., Kalthoff F., Huber N.E.,
 RA Baumruker T.;
 RT "A novel transiently expressed, integral membrane protein linked to
 RT cell activation. Molecular cloning via the rapid degradation signal
 RT AUUA.";
 RL J. Biol. Chem. 267:11267-11273(1992).
 CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
 CC NEUTRAL AMINO ACIDS. INVOLVED IN CELLULAR AMINO ACID UPTAKE.
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
 CC TRANSPORT PROTEIN SLC3A2/4F2HC.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN ADULT LUNG AND LIVER,
 CC AND IS ALSO EXPRESSED IN BRAIN, THYMUS, RETINA AND SOME OTHER
 CC TISSUES.
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
 CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

STRAIN=BALB/C;
MEDLINE=99115648; PubMed=9915839;
Nakamura E., Sato M., Yang H., Miyagawa F., Harasaki M., Tomita K.,
Matsuoka S., Noma A., Iwai K., Minato N.;
"4F2 (CD98) heavy chain is associated covalently with an amino acid
transporter and controls intracellular trafficking and membrane
topology of 4F2 heterodimer.";
J. Biol. Chem. 274:3009-3016(1999).
[2]
SEQUENCE FROM N.A.
STRAIN=ICR;
Kanai Y., Watanabe M., Endou H.;
"Localization of expression of system L neutral amino acid transporter
LAT1 in brain.";
Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
NEUTRAL AMINO ACIDS. INVOLVED IN CELLULAR AMINO ACID UPTAKE.
-!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
TRANSPORT PROTEIN SLC3A2/4F2HC.
-!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
-!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AB017189; BAA75520.1; -.
EMBL; AB023409; BAA90556.1; -.
MGD; MGI:1298205; Slc7a5.
InterPro; IPR002293; AA/rel_prmeasel.
InterPro; IPR004760; L_AA_transport.
InterPro; IPR004841; Permease.
Pfam; PF00324; aa_permeases; 1.
TIGRFAMS; TIGR00911; 2A0308; 1.
Transport; Amino-acid transport; Transmembrane.
TRANSMEM 51 71 POTENTIAL.
TRANSMEM 85 105 POTENTIAL.
TRANSMEM 121 141 POTENTIAL.
TRANSMEM 147 167 POTENTIAL.
TRANSMEM 174 194 POTENTIAL.
TRANSMEM 200 220 POTENTIAL.
TRANSMEM 248 268 POTENTIAL.
TRANSMEM 279 299 POTENTIAL.
TRANSMEM 324 344 POTENTIAL.
TRANSMEM 398 418 POTENTIAL.
TRANSMEM 436 456 POTENTIAL.
TRANSMEM 463 483 POTENTIAL.
FT CARBOHYD 50 50 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 8 8 R -> M (IN REF. 2).
SQ SEQUENCE 512 AA; 55902 MW; B3AE446E77374B0B CRC64;

Alignment Scores:
Pred. No.: 17.5 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x LAT1_MOUSE (1-512)

QY 517 TATGCTGAATTGGGAACAATA 540

Db 104 TyrAlaGluLeuGlyThrThrIle 111

RESULT 14

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF077866; AAC61479.1; -.
EMBL; AB018009; BAA84648.1; -.
EMBL; AF104032; AAD20464.1; -.
EMBL; AB018542; BAA33851.1; -.
EMBL; AB017908; BAA75746.1; -.
EMBL; M80244; AAA35780.1; -.
PIR; A42783; A42783.
Genew; HGNC:11063; SLC7A5.
MIM; 600182; -.
InterPro; IPR002293; AA/rel_prmeasel.
InterPro; IPR004760; L_AA_transport.
InterPro; IPR004841; Permease.
Pfam; PF00324; aa_permeases; 1.
TIGRFAMS; TIGR00911; 2A0308; 1.
Transport; Amino-acid transport; Transmembrane.
TRANSMEM 50 70 POTENTIAL.
TRANSMEM 84 104 POTENTIAL.
TRANSMEM 120 140 POTENTIAL.
TRANSMEM 146 166 POTENTIAL.
TRANSMEM 170 190 POTENTIAL.
TRANSMEM 199 219 POTENTIAL.
TRANSMEM 243 263 POTENTIAL.
TRANSMEM 274 294 POTENTIAL.
TRANSMEM 319 339 POTENTIAL.
TRANSMEM 396 416 POTENTIAL.
TRANSMEM 431 451 POTENTIAL.
TRANSMEM 458 478 POTENTIAL.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 15 15 A -> V (IN REF. 5).
FT CONFLICT 29 31 AKS -> SKR (IN REF. 5).
FT CONFLICT 35 35 S -> A (IN REF. 5).
FT CONFLICT 62 62 T -> A (IN REF. 5).
FT CONFLICT 88 88 V -> M (IN REF. 5).
FT CONFLICT 154 154 T -> A (IN REF. 5).
FT CONFLICT 230 230 N -> K (IN REF. 1).
SQ SEQUENCE 507 AA; 55010 MW; 767F3C60B62C0F02 CRC64;

Alignment Scores:
Pred. No.: 17.5 Length: 507
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x LAT1_HUMAN (1-507)

QY 517 TATGCTGAATTGGGAACAATA 540

Db 103 TyrAlaGluLeuGlyThrThrIle 110

RESULT 13

LAT1_MOUSE
ID LAT1_MOUSE STANDARD; PRT; 512 AA.

AC Q92127; Q9JMI4; Created

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Large neutral amino acids transporter small subunit 1 (L-type amino acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC).

GN SLC7A5 OR LAT1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

LAT1_RAT
ID LAT1_RAT STANDARD; PRT; 512 AA.
AC Q63016; Q9QWL4;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large neutral amino acids transporter small subunit 1 (L-type amino acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC) (Integral membrane protein E16) (TA1 protein).
DE SLC7A5 OR MPE16 OR TA1.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395066; PubMed=9726963;
RA Kanai Y., Segawa H., Miyamoto K., Uchino H., Takeda E., Endou H.;
RT "Expression cloning and characterization of a transporter for large neutral amino acids activated by the heavy chain of 4F2 antigen (CD98).";
RT J. Biol. Chem. 273:23629-23632(1998).
RN [2]
RP SEQUENCE OF 272-512 FROM N.A.
RC TISSUE=Hepatoma;
RX MEDLINE=95171385; PubMed=7532544;
RA Sang J., Lim Y.P., Panzica M., Finch P., Thompson N.L.;
RT "TA1, a highly conserved oncofetal complementary DNA from rat hepatoma, encodes an integral membrane protein associated with liver development, carcinogenesis, and cell activation.";
RT Cancer Res. 55:1152-1159(1995).
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE NEUTRAL AMINO ACIDS. INVOLVED IN CELLULAR AMINO ACID UPTAKE.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: EXPRESSED HEPATOMA BUT NOT IN NORMAL LIVER. ALSO EXPRESSED IN PLACENTA, TESTIS, BRAIN, OVARY, SPLEEN, MAMMARY GLAND, AND UTERUS.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT) SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB015432; BAA33035.1; -.
EMBL; U00995; AAA74411.1; -.
DR InterPro; IPR002293; AA/rel_prmease1.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 248 268 POTENTIAL.
FT TRANSMEM 279 299 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 398 418 POTENTIAL.
FT TRANSMEM 436 456 POTENTIAL.
FT TRANSMEM 463 483 POTENTIAL.
FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 512 AA; 55903 MW; B487CE0B58D73A02 CRC64;
Alignment Scores:
Pred. No.: 17.5 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0
US-09-667-170A-440 (1-2239) x LAT1_RAT (1-512)
QY 517 TATGCTGAATTGGGAACAATA 540
|||||
Db 104 TyrAlaGluLeuGlyThrThrIle 111
RESULT 15
LAT2_MOUSE
ID LAT2_MOUSE STANDARD; PRT; 531 AA.
AC Q9QXW9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large neutral amino acids transporter small subunit 2 (L-type amino acid transporter 2).
DE SLC7A8 OR LAT2.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=20044753; PubMed=10574970;
RA Rossier G., Meier C., Bauch C., Summa V., Sordat B., Verrey F., Kuhn L.C.;
RT "LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter of kidney and intestine.";
RT J. Biol. Chem. 274:34948-34954(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079165; PubMed=10610726;
RA Bassi M.T., Sperandeo M.P., Incerti B., Bulfone A., Pepe A., Surace E.M., Gattuso C., de Grandi A., Buoinconti A., Riboni M., Manzoni M., Andria G., Ballabio A., Borsani G., Sebastio G.;
RT "SLC7A8, a gene mapping within the lysinuric protein intolerance critical region, encodes a new member of the glycoprotein-associated amino acid transporter family.";
RT Genomics 62:297-303(1999).
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE NEUTRAL AMINO ACIDS. HAS HIGHER AFFINITY FOR L-PHENYLALANINE THAN LAT1. L-ALANINE IS TRANSPORTED AT PHYSIOLOGICAL CONCENTRATIONS.
CC PLAYS A ROLE IN BASOLATERAL (RE)ABSORPTION OF NEUTRAL AMINO ACIDS.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: MAINLY EXPRESSED IN KIDNEY AND INTESTINE.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT) SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF171668; AAF20380.1; -.
EMBL; Y19022; CAB69072.1; -.
DR MGD; MGI:1355323; Slc7a8.
DR InterPro; IPR002293; AA/rel_prmease1.

DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
FT TRANSMEM 309 329 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 387 407 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
FT TRANSMEM 446 466 POTENTIAL.
SQ SEQUENCE 531 AA: 57873 MW; AE9C3B42F3B24F8C CRC64;

Alignment Scores:
Pred. No.: 17.4 Length: 531
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x LAT2_MOUSE (1-531)

QY 1126 TCAATGCAGTGCAGTGACCTTT 1149

Db 298 SerAsnAlaValAlaValThrPhe 305
|||||

Search completed: April 16, 2003, 16:36:44
Job time : 33.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:27:06 ; Search time 96.5 Seconds
(without alignments)
9561.440 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 732
Sequence: 1 ggagggtgaagtgcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_091001_21199/app_query.fasta_1.2375
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=olin2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_122@runat_04042003_091001_21199 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	446	60.9	494	4	Q9BYH2	Q9byh2 homo sapien

2	57	7.8	92	4	Q9PLX0	Q9plx0 homo sapien
3	12	1.6	510	11	Q9QWS1	Q9qws1 mus musculu
4	12	1.6	510	11	Q9ZLK8	Q9zlk8 mus musculu
5	12	1.6	512	11	Q9R0S5	Q9r0s5 rattus norv
6	12	1.6	512	11	Q9QZ66	Q9qz66 rattus norv
7	10	1.4	523	4	Q9NS82	Q9ns82 homo sapien
8	10	1.4	530	11	Q9JMH8	Q9jmh8 mus musculu
9	9	1.3	340	8	Q9XMP1	Q9xmpl ceratitidis c
10	9	1.2	507	13	O57439	O57439 xenopus lae
11	9	1.2	507	13	O13020	O13020 xenopus lae
12	8	1.1	57	16	Q8RG74	Q8rg74 fusobacteri
13	8	1.1	65	11	Q8VED7	Q8ved7 mus musculu
14	8	1.1	82	16	Q9I5K5	Q9i5k5 pseudomonas
15	8	1.1	89	1	Q977T5	Q977t5 uncultured
16	8	1.1	100	16	Q930L9	Q930l9 rhizobium m
17	8	1.1	103	10	Q9C596	Q9c596 arabidopsis
18	8	1.1	104	3	Q05684	Q05684 saccharomyc
19	8	1.1	128	16	Q55384	Q55384 synechocyst
20	8	1.1	140	11	Q9D4S2	Q9d4s2 mus musculu
21	8	1.1	159	10	O65683	O65683 arabidopsis
22	8	1.1	160	10	P94060	P94060 arabidopsis
23	8	1.1	161	8	Q9TGJ4	Q9tgj4 taenia cras
24	8	1.1	180	7	Q9GIP4	Q9gip4 homo sapien
25	8	1.1	180	10	Q9SRR5	Q9srr5 arabidopsis
26	8	1.1	184	10	Q9S784	Q9s784 arabidopsis
27	8	1.1	189	5	Q9N8S7	Q9n8s7 trypanosoma
28	8	1.1	189	10	Q94K16	Q94k16 arabidopsis
29	8	1.1	203	4	Q96GR5	Q96gr5 homo sapien
30	8	1.1	217	5	Q24300	Q24300 drosophila
31	8	1.1	217	5	Q9V9D6	Q9v9d6 drosophila
32	8	1.1	241	16	Q8REU5	Q8reu5 fusobacteri
33	8	1.1	268	12	Q9EB06	Q9eb06 sesbania mo
34	8	1.1	272	11	Q99K91	Q99k91 mus musculu
35	8	1.1	278	16	Q9A101	Q9a101 streptococ
36	8	1.1	278	16	Q8XVM8	Q8xvm8 ralstonia s
37	8	1.1	281	17	Q8ZWH2	Q8zwh2 pyrobaculum
38	8	1.1	287	17	Q9UZG6	Q9uzg6 pyrococcus
39	8	1.1	288	17	Q8TSY3	Q8tsy3 methanosarc
40	8	1.1	292	5	Q94824	Q94824 tetrahymena
41	8	1.1	327	10	O64460	O64460 citrus jamb
42	8	1.1	327	10	O80421	O80421 citrus iyo.
43	8	1.1	327	10	Q9SAY9	Q9say9 citrus iyo.
44	8	1.1	327	10	Q9ZWH6	Q9zwh6 poncirus tr
45	8	1.1	327	10	Q8S988	Q8s988 microcitrus

ALIGNMENTS

RESULT 1
Q9BYH2
ID Q9BYH2 PRELIMINARY; PRT; 494 AA.
AC Q9BYH2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cystine/glutamate exchanger.
GN HXCT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.Y., Chairoungdua A., Cha S.H., Segawa H., Matsuo H., Kim D.K.,
RA Endou H., Kanai Y.,
RT "Human cystine/glutamate exchanger: cDNA cloning and upregulation by
RT oxidative stress in glioma cells."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB040875; BAB40574.1; -.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.

Db 56 ThrIleIleGlyAlaGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer 75

QY 454 GTGGGCGATGCTCTGACCATCTGGACGGTGTGTGGGCTCCTGTGCTACATATT 504
 |||||

Db 76 ValGlyMetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuphe 92

RESULT 3

Q9QWS1

ID Q9QWS1 PRELIMINARY; PRT; 510 AA.

AC Q9QWS1;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Y+LAT1b.

GN SLC7A7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NIH/SWISS; TISSUE=HEART;

RX MEDLINE=99094891; PubMed=9878049;

RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;

RT "Amino acid transport of y+L-type by heterodimers of 4F2hc/CD98 and

RT members of the glycoprotein-associated amino acid transporter

RT family.";

RL EMBO J. 18:49-57(1999).

DR EMBL; AJ130943; CAA10255.1; -.

DR MGD; MGI:1337120; SLC7A7.

DR InterPro; IPR002293; AA/rel_prmeasel.

DR InterPro; IPR004841; Permease.

DR Pfam; PF00324; aa_permeases; 1.

FT VARIANT 4 4 T -> S.

SQ SEQUENCE 510 AA; 55691 MW; 9F30FA7BAF126F6C CRC64;

Alignment Scores:

Pred. No.: 0.00362 Length: 510

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.64% Indels: 0

DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9QWS1 (1-510)

QY 517 TATGCTGAATGGGAACAACATATAAGAAATCTGGA 552
 |||||

Db 91 TyrAlaGluLeuGlyThrThrIleLysLysSerGly 102

RESULT 4

Q9Z1K8

ID Q9Z1K8 PRELIMINARY; PRT; 510 AA.

AC Q9Z1K8;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Y+LAT1a (Solute carrier family 7 (Cationic amino acid transporter, y+

DE system), member 7).

GN SLC7A7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=KIDNEY;

RX MEDLINE=99094891; PubMed=9878049;

RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;

RT "Amino acid transport of y+L-type by heterodimers of 4F2hc/CD98 and

RT members of the glycoprotein-associated amino acid transporter

RT family.";

RL EMBO J. 18:49-57(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ012754; CAA10170.1; -.

DR EMBL; BC014709; AAH14709.1; -.

DR MGD; MGI:1337120; SLC7A7.

DR InterPro; IPR002293; AA/rel_prmeasel.

DR InterPro; IPR004841; Permease.

DR Pfam; PF00324; aa_permeases; 1.

FT VARIANT 4 4 S -> T.

SQ SEQUENCE 510 AA; 55677 MW; 9F30FB1B88126F6C CRC64;

Alignment Scores:

Pred. No.: 0.00362 Length: 510

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.64% Indels: 0

DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9Z1K8 (1-510)

QY 517 TATGCTGAATGGGAACAACATATAAGAAATCTGGA 552
 |||||

Db 91 TyrAlaGluLeuGlyThrThrIleLysLysSerGly 102

RESULT 5

Q9R0S5

ID Q9R0S5 PRELIMINARY; PRT; 512 AA.

AC Q9R0S5;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Y+LAT1.

GN RY+LAT1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RA Kanai Y., Fukasawa Y., Segawa H., Endou H.;

RT "Characterization of a system y+L amino acid transporter.";

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB020520; BAA87325.1; -.

DR InterPro; IPR002293; AA/rel_prmeasel.

DR InterPro; IPR004841; Permease.

DR Pfam; PF00324; aa_permeases; 1.

SQ SEQUENCE 512 AA; 55641 MW; 708905B2ACB0A824 CRC64;

Alignment Scores:

Pred. No.: 0.00362 Length: 512

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.64% Indels: 0

DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9R0S5 (1-512)

QY 517 TATGCTGAATGGGAACAACATATAAGAAATCTGGA 552
 |||||

Db 93 TyrAlaGluLeuGlyThrThrIleLysLysSerGly 104

RESULT 6

Q9QZ66

ID Q9QZ66 PRELIMINARY; PRT; 512 AA.

AC Q9QZ66;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Amino acid transporter y+LAT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Ng A.M.L., Yao S.Y.M., Cheeseman C.I., Young J.D.;
RT "cDNA encoding rat jejunal amino acid transporter y+LAT1.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF200684; AAF07216.1; -.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
SQ SEQUENCE 512 AA; 55684 MW; 708905B2ACB2130B CRC64;

Alignment Scores:
Pred. No.: 0.00362 Length: 512
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9QZ66 (1-512)

QY 517 TATGCTGAATTGGGAACAACATAAAGAAATCTGGA 552
|||||
DB 93 TyrAlaGluLeuGlyThrThrIleLysSergly 104
|||||

RESULT 7

Q9NS82 ID Q9NS82 PRELIMINARY; PRT; 523 AA.

AC Q9NS82;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Asc-type amino acid transporter 1 (ASCI protein).
GN HASC-1 OR SLC7A10.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RX MEDLINE=20323545; PubMed=10863037;

RA Nakauchi J., Matsuo H., Kim D.K., Goto A., Chairoungdua A., Cha S.H.,

RA Inatomi J., Shiohara Y., Yamaguchi K., Saito I., Endou H., Kanai Y.;

RT "Cloning and characterization of a human brain Na+-independent

RT transporter for small neutral amino acids that transports D-serine

RT with high affinity.";

RL Neurosci. Lett. 287:231-235(2000).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21400982; PubMed=11509015;

RA Leclerc D., Wu Q., Ellis J.R., Goodyer P., Rozen R.;

RT "Is the slc7a10 gene on chromosome 19 a candidate locus for

RT cystinuria?";

RL Mol. Genet. Metab. 73:333-339(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RA Bassi M.T.;

RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RA Borsani G., Nunes V., Palacin M.;

RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB037670; BAB03213.1; -.

DR EMBL; AF340165; AAK93960.1; -.

DR EMBL; AF340155; AAK93960.1; JOINED.
DR EMBL; AF340156; AAK93960.1; JOINED.
DR EMBL; AF340157; AAK93960.1; JOINED.
DR EMBL; AF340158; AAK93960.1; JOINED.
DR EMBL; AF340159; AAK93960.1; JOINED.
DR EMBL; AF340160; AAK93960.1; JOINED.
DR EMBL; AF340161; AAK93960.1; JOINED.
DR EMBL; AF340162; AAK93960.1; JOINED.
DR EMBL; AF340163; AAK93960.1; JOINED.
DR EMBL; AF340164; AAK93960.1; JOINED.
DR EMBL; AJ277731; CAC81900.1; -.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
SQ SEQUENCE 523 AA; 56797 MW; 24BA0B36521AC2D4 CRC64;

Alignment Scores:
Pred. No.: 0.488 Length: 523
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.37% Indels: 0
DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9NS82 (1-523)

QY 409 GGAATCTTCATCTCTCTAAGGCGTGCTC 438
|||||

DB 57 GlyIlePheIleSerProLysGlyValleu 66
|||||

RESULT 8

Q9JMH8 ID Q9JMH8 PRELIMINARY; PRT; 530 AA.

AC Q9JMH8;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Asc1 (Asc-type amino acid transporter 1).

GN SLC7A10 OR ASC-1.

OS Mus musculus (Mouse), and

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090, 10116;

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES=Mouse;

RX MEDLINE=20200467; PubMed=10734121;

RA Fukasawa Y., Segawa H., Kim J.Y., Chairoungdua A., Kim D.K.,

RA Matsuo H., Cha S.H., Endou H., Kanai Y.;

RT "Identification and characterization of a Na+-independent neutral

RT amino acid transporter that associates with the 4F2 heavy chain and

RT exhibits substrate selectivity for small neutral d- and l- amino

RT acids.";

RL J. Biol. Chem. 275:9690-9698(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES=Rat; STRAIN=SPRAGUE DAWLEY; TISSUE=FOREBRAIN;

RA Alberati-Giani D., Kew J.N.;

RT "Cloning and characterization of a D-serine transporter from rat

RT brain.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB026688; BAA93617.1; -.

DR EMBL; AJ313387; CAC42885.1; -.

DR MGD; MGI:1858261; Slc7a10.

DR InterPro; IPR002293; AA/rel_prmeasel.

DR InterPro; IPR004841; Permease.

DR Pfam; PF00324; aa_permeases; 1.

SQ SEQUENCE 530 AA; 57548 MW; 0C5A80BF922DB54D CRC64;

Alignment Scores:

Pred. No.: 0.486 Length: 530

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.37% Indels: 0
DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9JMH8 (1-530)

QY 409 GGAATCTTCATCTCTCTCTAAGGCGTGCTC 438

Db 63 GlyIlePheIleSerProLysGlyValLeu 72

RESULT 9

Q9XMP1

ID Q9XMP1 PRELIMINARY; PRT; 340 AA.

AC Q9XMP1;

DT 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)

DE NADH dehydrogenase (ubiquinone) chain 2 (EC 1.6.5.3).

GN NDH-UL.

OS Ceratitis capitata (Mediterranean fruit fly).

OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Tephritoidea; Tephritidae; Ceratitis.

OX NCBI_TaxID=7213;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-20225575; PubMed=10762421;

RA Spanos L., Koutroumbas G., Kotsyfakis M., Louis C.;

RT "The complete sequence of the mitochondrial genome of the Medfly,

RT Ceratitis capitata.";

RL Insect Mol. Biol. 9:139-144(2000).

CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE - NAD(+) + UBIQUINOL.

DR EMBL; AJ242872; CAB45088.1; -

DR InterPro; IPR001750; Oxidored_q1.

DR Pfam; PF00361; oxidored_q1; 1.

KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.

SQ SEQUENCE 340 AA; 39231 MW; F4C646E1CE5E3F4F CRC64;

Alignment Scores:

Pred. No.: 6.2

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.26%

DB: 8

Length: 340

Matches: 9

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-667-170A-440 (1-2239) x Q9XMP1 (1-340)

QY 2196 CTCCTCTTTGCCATAATCATGATGAT 2170

Db 70 LeuLeuPheAlaIleMetMetTyr 78

RESULT 10

O57439

ID O57439 PRELIMINARY; PRT; 507 AA.

AC O57439;

DT 01-JUN-1998 (TReMBLrel. 06, Created)

DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE L amino acid transporter-1 LAT-1.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99047611; PubMed=9829974;

RA Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J.,

RA Shi Y.B., Zorzano A., Palacin M.;

RT "Identification and characterization of a membrane protein (y+L amino

RT acid transporter-1) that associates with 4F2hc to encode the amino
RT acid transport activity y+L. A candidate gene for lysinuric protein
RT intolerance.";

RL J. Biol. Chem. 273:32437-32445(1998).

DR EMBL; AF019906; AAB93541.1; -

DR InterPro; IPR002293; AA/rel_prmeasel.

DR InterPro; IPR004760; L_AA_transport.

DR InterPro; IPR004841; Permease.

DR Pfam; PF00324; aa_permeases; 1.

DR TIGRFAMS; TIGR00911; 2A0308; 1.

SQ SEQUENCE 507 AA; 55595 MW; EA0E9782B3313B82 CRC64;

Alignment Scores:

Pred. No.: 5.71

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.23%

DB: 13

Length: 507

Matches: 9

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-667-170A-440 (1-2239) x O57439 (1-507)

QY 610 TGGGTGGAACTCTCTATAATACGCCCT 636

Db 134 TrpValGluLeuLeuIleArgPro 142

RESULT 11

O13020

ID O13020 PRELIMINARY; PRT; 507 AA.

AC O13020;

DT 01-JUL-1997 (TReMBLrel. 04, Created)

DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Amino acid transporter chain (AmAt-L-1c) ASUR4.

GN ASUR4.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RX MEDLINE-97324256; PubMed=9178633;

RA Spindler B., Mastrobardino L., Custer M., Verrey F.;

RT "Characterization of early aldosterone-induced RNAs in A6 kidney

RT epithelia.";

RL Pflugers Arch. 434:323-331(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RX MEDLINE-98421678; PubMed=9751058;

RA Mastrobardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffing J.,

RA Shoemaker C.B., Verrey F.;

RT "Amino acid transport by heterodimers of 4F2hc/CD98 and members of a

RT new permease family.";

RL Nature 395:288-291(1998).

DR EMBL; Y12716; CAA73254.1; -

DR InterPro; IPR002293; AA/rel_prmeasel.

DR InterPro; IPR004760; L_AA_transport.

DR InterPro; IPR004841; Permease.

DR Pfam; PF00324; aa_permeases; 1.

DR TIGRFAMS; TIGR00911; 2A0308; 1.

FT CONFLICT 42 42 G -> A (IN REF. 0).

FT CONFLICT 199 199 A -> R (IN REF. 0).

FT CONFLICT 209 209 F -> L (IN REF. 0).

FT CONFLICT 391 391 N -> D (IN REF. 0).

SQ SEQUENCE 507 AA; 55529 MW; 2A6741CE89AB721D CRC64;

Alignment Scores:

Pred. No.: 5.71

Score: 9.00

Percent Similarity: 100.00%

Length: 507

Matches: 9

Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 13 Gaps: 0

US-09-667-170A-440 (1-2239) x O13020 (1-507)

QY 610 TGGTGGAACTCCTCATATACGCCCT 636

Db 134 TrpValGluLeuLeuLeuLeuLeuArgPro 142

RESULT 12

Q8RG74 PRELIMINARY; PRT; 57 AA.

AC Q8RG74;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein FN0443.
GN FN0443.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapattal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kypides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010556; AAL94639.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 57 AA; 6601 MW; 609693EE22C2F9AF CRC64;

Alignment Scores:
Pred. No.: 105 Length: 57
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 16 Gaps: 0

US-09-667-170A-440 (1-2239) x Q8RG74 (1-57)

QY 80 TTTCTTTTCTTCTTAATCTGTTT 57

Db 10 PheLeuPheSerLeuLeuLeuPhe 17

RESULT 13

Q8VED7 PRELIMINARY; PRT; 65 AA.

AC Q8VED7;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical 7.7 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019134; AAL19134.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 65 AA; 7700 MW; D2D9A9CC193308C5 CRC64;

Alignment Scores:
Pred. No.: 102 Length: 65
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q8VED7 (1-65)

QY 90 TTTTCTTCTTCTTCTTCTTCTTCTT 67

Db 31 phePhePheLeuSerPhePheLeu 38

RESULT 14

Q9I5K5 PRELIMINARY; PRT; 82 AA.

AC Q9I5K5;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Coat protein B of bacteriophage Pfl.
GN COAB OR PA0723.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004507; AAG04112.1; -.
DR HSSP; P03621; 1PFI.
KW Complete proteome.
SQ SEQUENCE 82 AA; 8377 MW; 25FD055DA4661504 CRC64;

Alignment Scores:
Pred. No.: 96.9 Length: 82
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 16 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9I5K5 (1-82)

QY 1432 CTGCAGTTGCTGGGCTGATTAT 1455

Db 69 LeuAlaValAlaGlyLeuIleTyr 76

RESULT 15

Q977T5 PRELIMINARY; PRT; 89 AA.

AC Q977T5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Small membrane protein.
OS uncultured crenarchaeote 4B7.
OC Archaea; Crenarchaeota; environmental samples;
OC marine archaeal group 1.
OX NCBI_TaxID=44557;
RN [1]
RP SEQUENCE FROM N.A.

RA Beja O., Koonin E.V., Aravind L., Taylor L.T., Seitz H., Stein J.L.,
RA Bensen D.C., Feldman R.A., Swanson R.V., DeLong E.F.,
RT "Comparative genomic analysis of coexisting archaeal genetic variants
in an Antarctic marine microbial assemblage,"
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40238; AAK66785.1; -.
SQ SEQUENCE 89 AA; 9325 MW; B4E02F76DBE2F393 CRC64;

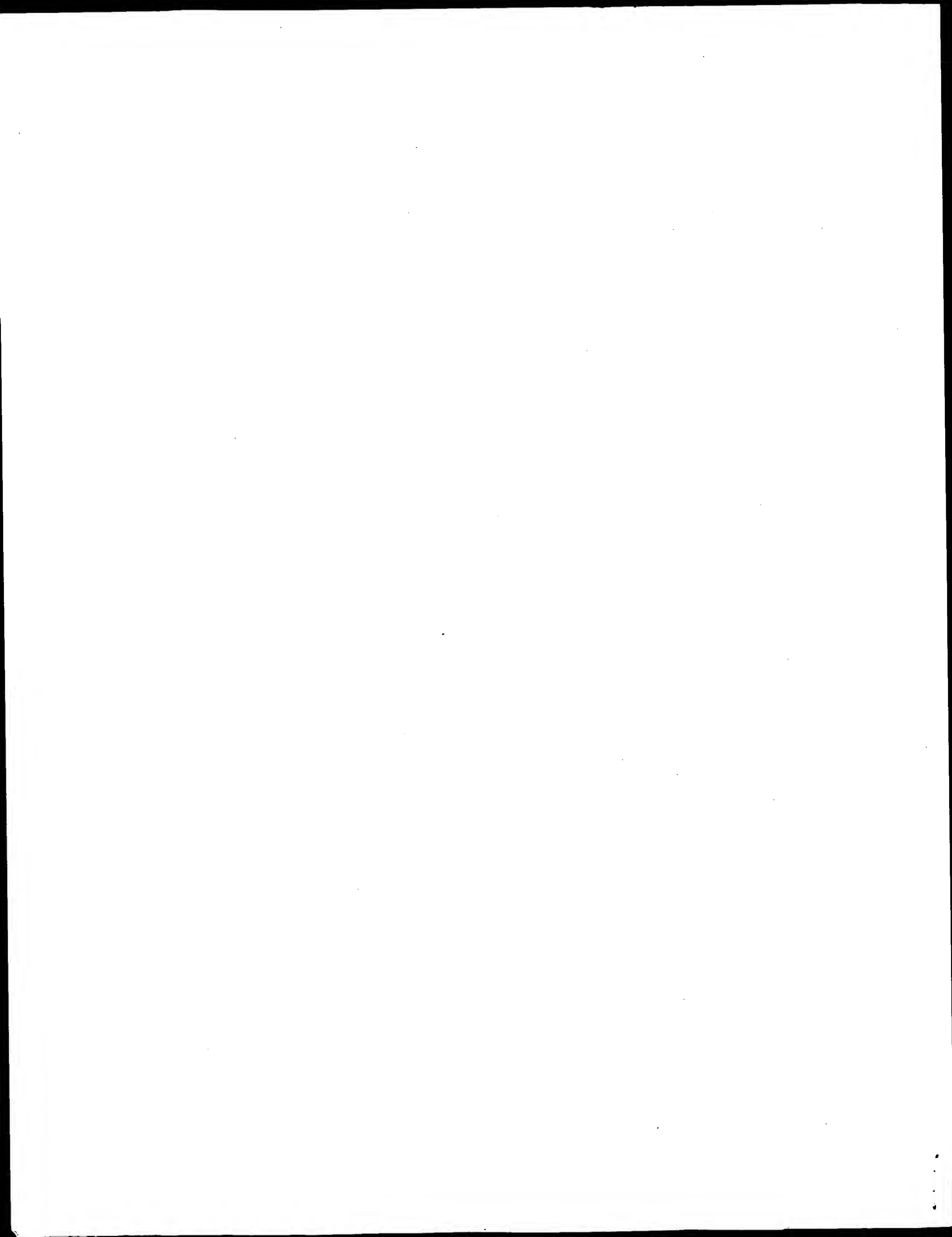
Alignment Scores:
Pred. No.: 95.3 Length: 89
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x Q977T5 (1-89)

QY 385 ATCATTTGGCACCACATCATTTGGAGCA 408

Db 19 IleIleGlyThrIleIleGlyAla 26

Search completed: April 16, 2003, 16:40:12
Job time : 104.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:06:50 ; Search time 50 Seconds
(without alignments)
8609.799 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/uspto_spool/us09667170/runat_04042003_090918_20307/app_query.fasta_1.2375
-DB=PIR_73 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_1_57@runat_04042003_090918_20307 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1210.5	29.4	507	JG0165	LAT1 protein - hum
2	1058.5	25.7	493	T21445	hypothetical prote
3	975	23.7	563	T32479	hypothetical prote
4	910	22.1	464	T28818	hypothetical prote
5	715	17.4	566	T15226	hypothetical prote
6	672.5	16.3	662	T32821	hypothetical prote
7	665.5	16.2	541	T31554	hypothetical prote
8	622.5	15.1	562	T16854	hypothetical prote
9	564	13.7	440	T24837	hypothetical prote
10	521.5	12.7	438	T69855	hypothetical prote
11	479.5	11.6	440	T89921	amino acid permeas
12	463	11.2	455	A11995	hypothetical prote
13	404.5	9.8	424	B69172	amino acid transpo
14	368.5	8.9	574	S61943	cationic amino aci
					methionine transpo

15	351.5	8.5	546	2	D87396	amino acid permeas
16	343.5	8.3	546	2	S48932	hypothetical prote
17	331	8.0	463	2	AE1155	amino acid transpo
18	331	8.0	583	2	T32266	hypothetical prote
19	330	8.0	503	2	T34694	probable cationic
20	328	8.0	463	2	AH1513	amino acid transpo
21	324	7.9	465	2	E69825	amino acid transpo
22	323	7.8	462	2	E65131	hypothetical 47.5
23	322	7.8	462	2	E91156	probable amino aci
24	322	7.8	462	2	B86002	probable amino aci
25	321.5	7.8	429	2	F71651	putrescine-ornithi
26	316.5	7.7	461	2	D69814	metabolite transpo
27	312.5	7.6	467	2	A97985	hypothetical prote
28	310.5	7.5	736	2	C69451	cationic amino aci
29	309.5	7.5	463	2	E95115	amino acid permeas
30	308.5	7.5	589	2	T32785	hypothetical prote
31	307.5	7.5	583	2	T48473	amino acid transpo
32	304	7.4	483	2	C82587	cationic amino aci
33	300.5	7.3	440	2	E70758	hypothetical prote
34	297	7.2	466	2	H97303	probable amino aci
35	295.5	7.2	427	2	F97790	putrescine-ornithi
36	295.5	7.2	441	2	C95307	probable transport
37	295	7.2	776	2	A84178	cationic amino aci
38	294.5	7.1	463	2	AH1758	amino acid transpo
39	292.5	7.1	463	2	AE1383	amino acid transpo
40	292	7.1	465	2	C97751	cationic amino aci
41	289.5	7.0	447	2	G70030	amino acid permeas
42	289	7.0	438	2	F97234	ethanolamin permea
43	287.5	7.0	470	2	F71686	cationic amino aci
44	286	6.9	590	2	A86307	amino acid transpo
45	282.5	6.9	585	2	E88087	protein B0454.6 [1

ALIGNMENTS

RESULT 1
JG0165

LAT1 protein - human

C:Species: Homo sapiens (man)

C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000

C:Accession: JG0165; A42783

R:D.Prasad, P.; Wang, H.; Huang, W.; Kekuda, R.; P.Rajan, D.; H.Leibach, F.; Ganapath

Biochem. Biophys. Res. Commun. 255, 283-288, 1999

A:Title: Human LAT1, a subunit of system L amino acid transporter: Molecular cloning

A:Reference number: JG0165; MUID:99160855; PMID:10049700

A:Accession: JG0165

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-507 <DAP>

A:Cross-references: GB:AF104032; NID:g4426639; PIDN:AAD20464.1; PID:g4426640

R:Gaugitsch, H.W.; Prieschl, E.E.; Kalthoff, F.; Huber, N.E.; Baumruker, T.

J. Biol. Chem. 267, 11267-11273, 1992

A:Title: A novel transiently expressed, integral membrane protein linked to cell acti

A:Reference number: A42783; MUID:92283834; PMID:1597461

A:Accession: A42783

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 267-507 <GAU>

A:Cross-references: GB:M80244; NID:g181907; PIDN:AAA35780.1; PID:g181908

A:Experimental source: peripheral blood lymphocytes

C>Note: sequence extracted from NCBI backbone (NCBIN:104749, NCBI:104750)

C:Superfamily: arginine permease

C:Keywords: transmembrane protein

Alignment Scores:

Pred. No.: 6.65e-100 Length: 507
Score: 1210.50 Matches: 232
Percent Similarity: 66.19% Conservative: 95
Best Local Similarity: 46.96% Mismatches: 164
Query Match: 29.38% Indels: 3
DB: 2 Gaps: 2

US-09-667-170A-440 (1-2239) x JG0165 (1-507)

Db 88 LeuGlyThrLeuIleLysSerGlyGlyAspTyrAlaTyrIleMetGluAlaPheGly 107
QY 586 CCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAATACGCCCTGCAGCTACT 645
Db 108 PropheValAlaPheIleArgLeuTrpIleGluAlaIleValValArgProCysThrVal 127
QY 646 GCTGTGATATCCCTGGCAFTTGGACGCTACATTTCTGGAACCAATTTTATTCAATGTGAA 705
Db 128 ThrIleValAlaLeuThrPheAlaIleTyrGlyLeuArgProPhePheProAspCysAla 147
QY 706 ATCCCTGAACTTGGATCAAGCTCATACAGCTGTGGCATACTGTAGTGATGGTCTTA 765
Db 148 ProProAspValValAlaGluLeuLeuAlaIleLeuLeuIleValLeuMetThrAlaIle 167
QY 766 AATAGCATGAGTGCAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTC 825
Db 168 AsnCysIleSerValArgLeuAlaThrIleValGlnAspTrpPheThrIleAlaIysVal 187
QY 826 ACAGCAATCTCTGATAAATATAGTCCCTCGAGTGTATGCAGCTAATTAAGGTCAAACGCAG 885
Db 188 ValAlaLeuCysIleIleLeuThrGlyLeuGlyLeuLeuPhePheGlyGluSerGln 207
QY 886 -----AACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGCGGTGCCA 936
Db 208 TyrLysAspSerPheGluAsnIlePheGluAsnThrSerGlnAspPheThrLysValSer 227
QY 937 CTGGCTTTTATATGGAATGTATGCATATGCTGGCTGGTGTACCTCAACATTTGTACT 996
Db 228 LeuAlaPheTyrSerGlyLeuPheAlaTyrSerGlyTrpAsnPheLeuAsnPheIleVal 247
QY 997 GAAGAAGTAGAAACCCCTGAAAAACCACTTCCCTTGCATATGTATATCCATGCCATT 1056
Db 248 GluGluLeuGlnAsnProLysArgAsnLeuProLeuAlaIleAlaIleSerIleThrSer 267
QY 1057 GTCACCATTTGGCTATGTGTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAGGAG 1116
Db 268 CysThrValIleTyrValLeuThrAsnValAlaLeuTyrThrAlaIleSerProAspGlu 287
QY 1117 CTGCTGCTTCAATGACAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCA 1176
Db 288 MetLeuGluSerProAlaValAlaValLeuPheAlaAsnLysLeuTyrGlyLysPheAla 307
QY 1177 TTAGCAGTTCGGATCTTTGTGCTCCTCTCCTGCTGGCTCCATGAACGGTGGTGT 1236
Db 308 PheIleMetProLeuCysValAlaCysSerThrIleGlySerAlaAsnGlyValIlePhe 327
QY 1237 GCTGTCTCCAGGTTATTTCTATGTTGGCTCTCCGAGGGTCCACCTTCCAGAAATCCTCTCC 1296
Db 328 ThrSerAlaArgLeuPheTyrSerGlyAlaArgGluGlyGlnMetProAlaValLeuThr 347
QY 1297 ATGATTCATGTCGCGAAGCACACTCCTCTACAGCTGTTTATTTGTCACCCCTTTGACA 1356
Db 348 MetIleAsnLysLysThrLysThrProIleProAlaValIleLeuThrGlyAlaLeuSer 367
QY 1357 ATG---ATAATGCTCTTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCTCAGTTTGCC 1413
Db 368 IleAlaTyrLeuLeuAlaSerLysAspValTyrGlnLeuIleAsnTyrIleGlnIleSer 387
QY 1414 AGGTGGCTTTTATTTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 388 TyrTrpLeuAlaIleGlyThrAlaIleAlaAlaLeuPheTrpLeuArgArgThrMetPro 407
QY 1474 GATATGCATCGTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATCATGTC 1533
Db 408 AspAlaSerArgProIleLysValProLeuIleTrpProAlaIlePheLeuAlaGlyCys 427
QY 1534 CTCTTCATGGTGGCTTTCCCTCTATTCGGACCCCAATTTAGTACAGGATTTGGCTTCGTC 1593
Db 428 IleAlaLeuValLeuIleProMetValAlaAlaProArgAspThrGlyIleGlyLeuLeu 447
QY 1594 ATCACTCTGAGTGGAGTCCCTGGCTATTATCTCTTTTATATATGGGACAAACCCAGG 1653
Db 448 IleMetLeuSerAlaValProValTyrGlyIlePheIleGlyTrpLysAsnLysProLys 467

QY 1654 TGGTTTGAATAATGTACAGAGAAAATAACAGAACATTACAAATAATACTGGAAGTTGTA 1713
Db 468 TrpPheAsnGluPheIleAspSerSerThrValPheIleGlnLysLeuPheMetValVal 487
QY 1714 CCAGAAGAAGATAAG 1728
Db 488 ---AspGluAspLys 491
RESULT 3
T32479
hypoetical protein F52H2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Jun-2000
C:Accession: T32479
R:Chisoe, S.; Hawkins, J.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid F52H2.
A:Reference number: 221175
A:Accession: T32479
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-563 <CHI>
A:Cross-references: EMBL:AF026214; PIDN:AAB71312.1; GSPDB:GN00028; CESP:F52H2.2
A:Experimental source: strain Bristol N2; clone F52H2
C:Genetics:
A:Gene: CESP:F52H2.2
A:Map position: X
A:Introns: 20/3; 60/1; 123/2; 162/1; 204/1; 255/2; 296/3; 329/1; 401/2; 433/3; 544/1
C:Superfamily: arginine permease
Alignment Scores:
Pred. No.: 8.36e-79 Length: 563
Score: 975.00 Matches: 190
Percent Similarity: 55.85% Conservative: 106
Best Local Similarity: 35.85% Mismatches: 160
Query Match: 23.67% Indels: 74
DB: 2 Gaps: 3
US-09-667-170A-440 (1-2239) x T32479 (1-563)

QY 346 CTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGA 405
Db 28 LeuGluLysThrMetThrLeuPheAsnGlyValSerIleIleValGlyCysIleIleGly 47
QY 406 GCAGGAATCTTCATCTCTCTAAGGGCGTGTCTCCAGAACACGGCGGCGCATGTCT 465
Db 48 SerGlyIlePheIleSerProThrGlyIleGlnAlaGlnAlaGlySerValGlyLeuSer 67
QY 466 CTGACCATCTGGACGGTGTGTGGGTCTGTCTCATTATTTGGAGCTTTGTCTTATGCTGAA 525
Db 68 LeuIleValTrpValLeuSerGlyLeuPheAlaGlyIleGlyAlaPheCysTyrAlaGlu 87
QY 526 TTGGGAACAACATAAAGAAATCTGGAGGTCACTTACACATATATTTTGGAAAGTCTTTGGT 585
Db 88 LeuGlyThrLeuIleArgLysSerGlyGlyAspTyrAlaTyrIleMetGluAlaPheGly 107
QY 586 CCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATAATACGCCCTGCAGCTACT 645
Db 108 ProPheLeuAlaPheLeuArgLeuTrpIleGluSerIleValValArgProCysThrAla 127
QY 646 GCTGTGATATCCCTGGCAATTTGGACGGTACATTTCTGGAACCACTTTTATTCATGTA 705
Db 128 ThrIleValAlaLeuThrPheAlaIleTyrMetLeuLysProPheTyrProAspCysAsp 147
QY 706 ATCCCTGAACCTTGGATCAAGCTCATACAGCTGTGGCATAACTGTAGTGATGGTCTTA 765
Db 148 SerProProLeuSerThrGluLeuIleAlaAlaLeuLeuLeuValLeuLeuThrAlaVal 167
QY 766 AATAGCATGAGTGTACGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTC 825
Db 168 AsnCysIleSerValLysTrpAlaSerLysValGlnAspPhePheValThrLysThr 187

QY 826 ACAGCAATCTGATAATATATAGTCCCTGGAGTTATGACGCTAATTAAGGTCAACGCGAG 885
Db 188 AlaAlaLeuValleuIlePheThrGlyLeuTrpAsnMetValSerGlyLysProGlu 207
QY 886 -----AACTTTAAAGAGCGCTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCA 936
Db 208 AlaPheAspSerPheGluAsnIlePheGluAsnThrAlaLysAspLeuGluThrAlaSer 227
QY 937 CTGGCTTTTATTATGGAATGATGTCATATGCTGGCTGGTTTACTCAACTTTTGTACT 996
Db 228 LeuAlaPheTyrSerGlyLeuPheAlaTyrGlnGlyTrpAsnTyrLeuAsnPheIleVal 247
QY 997 GAAGAGTAGAAACCCCTGAAACCAATTCCTCCCTTGCAATATGATATATCCATGGCCATT 1056
Db 248 GluGluLeuGlnAsnProLysArgAsnLeuProLeuSerIleAlaIleSerCysSerLeu 267
QY 1057 GTCACATTGGCTATGTCGTGACAAATGTGGCTACTTTTACGACCAATTAATGCTGAGGAG 1116
Db 268 CysThrIleIleTyrThrLeuThrAsnValAlaLeuTyrThrSerIleThrProAspGlu 287
QY 1117 CTGCTGCTTCAATGACGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCA 1176
Db 288 MetLeuAlaSerProAlaValAlaValLeuPheAlaGluLysAsnTyrGlyTrpPheAla 307
QY 1177 TTAGCAGTCCGATCTTTGTTGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTG--- 1233
Db 308 PheCysMetProIlePheValAlaCysSerThrIleGlySerAlaAsnGlyValIleLeu 327
QY 1233 ----- 1233
Db 328 ThrArgLeuCysGlySerArgAspValArgGlnAlaAlaIleAlaGluValAlaGluPro 347
QY 1233 ----- 1233
Db 348 GlnAsnProLysIlePheGlyAsnArgGlnLeuProAsnLeuThrSerGlnLysIlePhe 367
QY 1233 ----- 1233
Db 368 GlyAsnArgGlnLeuProPhePheGluAsnPhePheValGluIleIlePheLeuSerAla 387
QY 1234 -----TTTGCTGTCTCCAGTTATTCTATGTTGGTCT 1266
Db 388 LeuGluLysSerMetValValPhePheTyrSerGluThrLeuPheTyrCysGlyAla 407
QY 1267 CGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTTCATGTCGCGAAGCACACTCTCTA 1326
Db 408 ArgGluGlyGlnMetProAsnValLeuThrMetValAsnLysThrThrLysThrProIle 427
QY 1327 CCAGCTGTTATTGTTTGCACCCCTTTGACAAATGATA---ATGCTCTTCTCTGGAGACCTC 1383
Db 428 ProAlaValIleLeuThrGlyLeuLeuSerLeuLeuTyrLeuLeuLeuSerAsnIle 447
QY 1384 GACAGTCTTTTGAATTTCTCAGTTTGGCCAGGTGGCTTTTATTTGGCTGGCAGTTGCT 1443
Db 448 TyrSerLeuIleAsnTyrIleGlnValSerTyrTrpIleAlaIleGlyGlyAlaIleLeu 467
QY 1444 GGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCCTTTCAAGGTGCCACTG 1503
Db 468 AlaLeuPheTyrPheArgLysThrMetProAspAlaProArgAlaValLysAlaProIle 487
QY 1504 TTCATCCAGCTTTGTTTCTTCCATGCTTCATGTTGGCTTTCCTTCCCTCTATTCTG 1563
Db 488 ValPheProIleIlePhePheIleGlyCysValLeuLeuValLeuValProValLeuGly 507
QY 1564 GACCCATTATTAGTACAGGATTGGCTTCTGTCATCTGCTGAGTCCCTGCGTATTAT 1623
Db 508 AsnProLysAspThrAlaIleGlyIleLeuIleMetLeuSerGlyValProValTyrLeu 527
QY 1624 CTCTTTATTATATGGACAAAGAACCCAGGTGGTTTAGAATAATGTACAGAAATAACC 1683
Db 528 IlePheIleAlaTrpLysGlyLysProLysCysIleAspSerLeuThrAspSerValThr 547
QY 1684 AGAACATTACAAATAATAACTGGAAGTTGTA 1713

Db 548 IlePheThrGlnLysLeuPheMetValVal 557
RESULT 4
T28818
hypothetical protein F07C3.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C;Accession: T28818
R;Favella, A.; Gattung, S.
submitted to the EMBL Data Library, March 1996
A;Description: The sequence of C. elegans cosmid F07C3.
A;Reference number: Z20528
A;Accession: T28818
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-464 <FAV>
A;Cross-references: EMBL:U50308; PIDN:AAC48006.1; GSPDB:GN00023; CESP:F07C3.7
A;Experimental source: strain Bristol N2; clone F07C3
C;Genetics:
A;Gene: CESP:F07C3.7
A;Map position: 5
A;Introns: 91/3; 240/2; 272/3; 299/2
C;Superfamily: arginine permease
Alignment Scores:
Pred. No.: 5.37e-73 Length: 464
Score: 910.00 Matches: 185
Percent Similarity: 60.81% Conservative: 99
Best Local Similarity: 39.61% Mismatches: 159
Query Match: 22.09% Indels: 24
Gaps: 5
US-09-667-170A-440 (1-2239) x T28818 (1-464)
QY 337 AAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATATCATTTGGCACC 396
Db 12 GlnIleLysLeuLysProArgIleSerLeuPheAsnGlyCysThrIleIleGlyVal 31
QY 397 ATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTGTCTCCAGAACACGGGCGGTG 456
Db 32 IleIleGlySerGlyIlePheValSerProLysGlyValLeuLeuGluAlaGlySerAla 51
QY 457 GGCATGCTCTGACCATCTGGACGGTGTGTGGGCTCTGTCTACATTTGGAGCTTTGTCT 516
Db 52 GlyMetSerLeuLeuIleTrpLeuLeuSerGlyValPheAlaMetIleGlyAlaValCys 71
QY 517 TATGCTGAATTTGGGAACAACATAAAGAAATCTGAGGCTCATACACATATATTTGGAA 576
Db 72 TyrSerGluLeuGlyThrLeuIleProLysSerGlyGlyAspTyrAlaTyrIleTyrGlu 91
QY 577 GTCTTTGCTCCATTACAGCTTTTGTACGAGTCTGSGTGGAACTCTCTCAATAACGCCCT 636
Db 92 AlaPheGlyProLeuProSerPheLeuPheLeuTrpValAlaLeuValIleIleAsnPro 111
QY 637 GCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTTATT 696
Db 112 ThrSerLeuAlaIleIleAlaIleThrCysAlaThrTyrAlaLeuGlnProPheTyr--- 130
QY 697 CAATGTGAATCCCTGAACCTTGCATCAAGCTCATTTACAGCTGTGGCATACTGTAGTG 756
Db 131 SerCysProValProAspValValValAsnLeuPheAlaGlyCysIleIleAlaValLeu 150
QY 757 ATGGTCTCTAAATAGCATGAGTGTGCTGAGCGCGCGGATCCAGATTTTCTTAACCTTT 816
Db 151 ThrPheIleAsnCysTrpAspValArgMetAlaThrArgThrAsnAspPhePheThrIle 170
QY 817 TGCAAGCTCACAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGCAGCTAATTAAGGT 876
Db 171 ThrLysLeuIleAlaLeuThrLeuIleIleThrCysGlyGlyTyrTrpLeuSerLeuGly 190
QY 877 CAAACGCAGAAC-----TTTAAAGACGCTTTTTCAGGAAGAGATTCAGTATTACGCGG 930

Db 191 HisIleAspAsnLeuValMetProaspValAlaGluGlySerGlnThrLysLeuSerAla 210
 QY 931 TTGCCACATGGCTTTTATTATGAATGATGATATGCTGGCTGGTTTACCTCAACTTT 990
 Db 211 ileAlaMetAlaPheTyrSerGlyValPheSerPheSerGlyPheSerTyrLeuAsnPhe 230
 QY 991 GTTACTGAAGAAGTAGAAACCCCTGAAAAACCAATCCCTTGGCAATATGATATCCATG 1050
 Db 231 ValThrGluGluLeuLysAsnProPheArgAsnLeuProArgAlaIleTyrIleSerIle 250
 QY 1051 GCCATGTCCACCATGGCTATGCTGACAAATGTCCTTACCTTACGACCATTAATGCT 1110
 Db 251 ProIleValThrIleValTyrMetLeuValAsnIleAlaTyrPheSerValLeuThrVal 270
 QY 1111 GAGGAGCTGCTGCTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAAT 1170
 Db 271 AspGluIleLeuAspSerAspAlaValAlaIleThrPheAlaAspLysIleLeuGlyThr 290
 QY 1171 TTCTCATTAGCAGTTCGATCTTGTGGCTCTCTCTGCTTGGCTCCATGAACGGTGGT 1230
 Db 291 Phe-----GlySer 293
 QY 1231 GTGTTTGCTGCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATC 1290
 Db 294 LysIleLeuMetProLeuMetPhePheSerGlyAlaArgAsnSerGlnLeuProGluLeu 313
 QY 1291 CTCTCCATGATTCATGTCGCGCAAGCACACTCTCTACAGTGTTTATGTTTGGCACCCT 1350
 Db 314 PheAlaMetIleSerIleArgGlnLeuThrProIleProSerLeuIlePheLeuGlyGly 333
 QY 1351 TTGACAATGATAATGCTCTTCTCGGAGACCTCGACAGTCTTTTGAATTTCTCTCAGTTT 1410
 Db 334 ThrSerIleValMetLeuPheIleGlyAsnValPheGlnLeuIleAsnTyrLeuSerPhe 353
 QY 1411 GCCAGGTGGCTTTTATTGGGTGGCAGTGTGCTGGGCTGATTTATCTTCGATACAAATGC 1470
 Db 354 AlaGluSerLeuValPheSerSerValAlaGlyLeuLeuLysLeuArgPheThrMet 373
 QY 1471 CCAGAT-----ATGCATCGTCTCTTCAAGTGGCCACTGTTTCATCCAGCTTTGTTT 1521
 Db 374 ProGluAsnValLeuAsnAlaArgProIleLysIleSerLeuLeuTrpProIleLeuPhe 393
 QY 1522 TCCTTCACATGCCTCTTCATGTTGCCCTTCCCTCTAT--TCGGACCATTTAGTACA 1578
 Db 394 PheLeuMetCysLeuPheLeuIleLeuProPhePheHisSerAspProTrpGluLeu 413
 QY 1579 GGGATTGGCTTCGTATCACTCTGACTGGAGTCCCTCGGTATTTATCTCTTATATATGG 1638
 Db 414 IleTyrGlyValPheLeuValLeuSerGlyIleProIleTyrValLeuPheValTyrAsn 433
 QY 1639 GACAAGAAACCCAGGTGGTTAGATAATGTCAGAGAAAATAACCAAGAACATTTACAATA 1698
 Db 434 LysTyrArgProGlyPheIleGlnSerValTrpIleGlyPheThrHisPheIleGlnLys 453
 QY 1699 ATACTGGAAGTTGTACCAGAA 1719
 Db 454 LeuPheTyrCysValProGlu 460

RESULT 5

T15226
 hypothetical protein C55C2.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15226
 R:Goela, D.
 submitted to the EMBL Data Library, May 1997
 A:Description: The sequence of C. elegans cosmid C55C2.
 A:Reference number: Z18311
 A:Accession: T15226
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-566 <GOE>
 A:Cross-references: EMBL:AF003144; NID:g2088754; PID:g2088757; PIDN:AB54195.1; GSPDB:GN

A:Experimental source: strain Bristol N2; clone C55C2

C:Genetics:

A:Gene: CESP:C55C2.5

A:Map position: 1

A:Introns: 85/1; 148/3; 170/1; 214/1; 303/3; 370/3; 426/2; 481/1; 550/2

Alignment Scores:

Pred. No.: 1.62e-55 Length: 566
 Score: 715.00 Matches: 159
 Percent Similarity: 54.85% Conservative: 107
 Best Local Similarity: 32.78% Mismatches: 205
 Query Match: 17.35% Indels: 14
 DB: 2 Gaps: 6

US-09-667-170A-440 (1-2239) x T15226 (1-566)

QY 334 GAGAAAGTGCAGCTG-----AAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATT 384
 Db 25 GluAsnThrGlnIleMetAspGluSerHisLysMetGlyPheLeuGlyAlaThrSerTyr 44
 QY 385 ATCATTGGCACCACATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGTGTCCAGAAC 444
 Db 45 ValIleGlyAsnIleIleGlySerGlyIlePheIleThrProAlaSerIleLeuArgAsn 64
 QY 445 ACGGCGACGCTGGGCATGTCTCTGACCATCTGGACGGTGTGGGGTCTCTGTCTCACTATT 504
 Db 65 ValAspSerIleGlyLeuSerLeuIleTrpValLeuCysAlaValIleAlaIleLeu 84
 QY 505 GGAGCTTTGTCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCACTACACA 564
 Db 85 GlyAlaIleCysTyrIleGluLeuGlyThrSerIleArgGluAlaGlyCysAspPheAla 104
 QY 565 TATATTTTGAAGTCTTTGGTCCATTACCAGTCTTGTACGAGTCTGGGTGGAACTCCTC 624
 Db 105 TyrIleCysTyrValLysTrpTyrSerIleAlaPheAlaPheMetTrpValSerValLeu 124
 QY 625 ATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGCTACATTTCTGGAA 684
 Db 125 MetThrTyrProAlaThrIleAlaIleCysAlaGluThrPheGlyGlnTyrLeuIleGlu 144
 QY 685 CCATTTTTCATATGTAATGTAATCCCTGAACTTCCGATC-----AAGCTCAT 732
 Db 145 GlyLeuLysGlnTyrTyrGluIleAspAspAlaLeuValProThrCysGlnLysLeuPhe 164
 QY 733 ACAGCTGTGGGCATAACTGTAGTGTGTCCTTAAATAGCATGAGTGTGAGC---TGGAGC 789
 Db 165 AlaTyrSerLeuLeuPheLeuValThrTrpMetAsnPhePheGluLeuSerLysPheAla 184
 QY 790 GCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAATATAGTC 849
 Db 185 AlaArgPheGlnIleLeuAlaThrIleAlaLysLeuPheSerCysMetLeuIleIleGly 204
 QY 850 CCTGGA---GTTATGCAGCTAATTAAGGTCAAACCGCAGAACCTTTAAAGACGCCCTTTTCA 906
 Db 205 ThrGlyPheTyrPheTyrPheValLysGlyTyrHisGlyTyrLeuGluAsnProMetGln 224
 QY 907 GGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATCATAT 966
 Db 225 GlySerLysTyrGlyThrGlyAsnLeuIleLeuGlyPheTyrGlyGlyLeuTrpAlaPhe 244
 QY 967 GCTGGCTGGTTTACCTCAACTTTGTTACTGAGAAGTAGAAACCCCTGAAAAAACCAT 1026
 Db 245 SerGlyTrpAspValLeuAsnTyrSerThrGlyGluIleLysHisProLysArgAsnVal 264
 QY 1027 CCCTTGTCAATATGATATCCATCCATGGCCATTTGTCACCATTGGCTATGTGCTGACAAATGTG 1086
 Db 265 PropheAlaLeuLeuThrGlyIleSerValValThrAlaIleTyrValAlaIleAsnVal 284
 QY 1087 GCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAATGCAGTGGCAGTGACC 1146
 Db 285 AlaTyrPheValValLeuAspValGluThrValLysGlnSerAspAlaValAlaIle 304
 QY 1147 TTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCC 1206

Thu Apr 17 09:48:26 2003

||||| 305 PheSerArgGluThrLeuGlyAspPheAlaAsnValValPropheLeuIleGlyIleLeu 324
1207 TGCTTTGGCTCCATGAACGGTGGTGTGTTGGTGTCTCCAGGTATCTATGTTGGCTCT 1266
325 LeuIleGlySerLeuAsnSerLeuPheSerGlySerArgTyrMetTyrAlaAlaAla 344
1267 CGAGAGGCTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGACACACTCCTCTA 1326
345 ArgGlnGlyHisLeuProAlaCysPheSerCysValAsnThrGluThrGluSerProArg 364
1327 CCAGCTGTATTTGTTTGCACCTTTTGACAAATGATAATGCTCTCTCTGGAGACCTCGAC 1386
365 ValAlaValLeuAlaGlnSerValLeuAlaLeuValIleSerTyrIleGlyAspLeuAsp 384
1387 AGTCTTTTGAATTTCTCCAGTTTTCAGGTTGCTTTTATTTGGCTGGCAGTTGCTGGG 1446
385 ThrLeuIleThrTyrValMetPheGlyPheTrpAlaGlnArgIlePheSerIleValAla 404
1447 CTGATTTATCTCGATACAAATGCCAGATATGCAAT---CGTCTTTTCAAGGTGCCACTG 1503
405 LeuLeuIleIleArgHisAsnHisIleProValHisProAspAlaValArgValProLeu 424
1504 TTCATCCCAGCTTTGTTTTCCTTCACATGCTCTTTCATGGTGGCTTTCCTCTATTCG 1563
425 PheCysIleTyrLeuPheLeuAlaIleThrValAlaLeuValIleIleProIlePheTyr 444
1564 GACCCATTAGTACAGGATTGGCTTCGTCATCACTCTGCTGAGTCCCTGCGTATTAT 1623
445 GluPheGlnSerThrAlaLeuAlaIleAlaIleCysLeuPheGlyPheValLeuTyrTyr 464
1624 CTCTTTTATATATGGACAAAGAAACCCAGGTGGTTTAGAATAATGTACAGAAATAAACC 1683
465 ValPheIleHisLysAlaIlePheProArgTrpLeuValAlaLeuAsnLysLysValThr 484
1684 AGAACATTACAAATAATACTGGAAGTTGTACCAAGAAAGATAGTATGAACTAATGGA 1743
485 LeuTrpCysCysIleLeuPheAspCysLeuProAspVal-----LysGlyGly 500
1744 CTTGAGATCTTGGCA 1758
501 ValGlnLeuLeuAla 505
RESULT 6
T32821
hypoetical protein F54D12.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32821
R:Maggi, L.; Goela, D.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid F54D12.
A:Reference number: 221228
A:Accession: T32821
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-662 <MAG>
A:Cross-references: EMBL:AF040647; PIDN:AAB94996.1; GSPDB:GN00020; CESP:F54D12.3
A:Experimental source: strain Bristol N2; clone F54D12
C:Genetics:
A:Gene: CESP:F54D12.3
A:Map position: 2
A:Introns: 31/1; 88/2; 191/1; 235/1; 304/1; 396/2; 443/3; 510/3; 584/3
Alignment Scores:
Pred. No.: 1,06e-51 Length: 662
Score: 672.50 Matches: 178
Percent Similarity: 43.01% Conservative: 96
Best Local Similarity: 27.94% Mismatches: 186
Query Match: 16.32% Indels: 177
DB: 2 Gaps: 13

US-09-667-170A-440 (1-2239) x T32821 (1-662)

QY 337 AAAGTCGAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGGTCTCCATTATCATTTGGCACC 396
Db 7 GluIleProGlnLysHisThrIleGlyLeuIleThrAlaIleSerTyrThrValGlyAsp 26
QY 397 ATCATTTGAGCAGGAATCTTCATCTCTCTAAGGGCGTGTCTCCAGAACACGGCAGCGTG 456
Db 27 IleValGlySerGlyIlePheIleSerProThrSerIleLeuAsnHisAlaGlySerVal 46
QY 457 GGCATGTCTGTGACCATCTGGACGGTGTGTGGGTCTCTGCTCATTTTGGAGCTTGTCT 516
Db 47 GlyLeuSerLeuCysLeuTrpAlaLeuCysAlaCysIleSerLeuPheGlyAlaLeuSer 66
QY 517 TATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCTTACACATATATTT----- 570
Db 67 TyrValGluLeuGlyThrSerIleArgLysSerGlyCysAspPheAlaTyrLeuSerHis 86
QY 570 ----- 570
Db 87 PheGlySerThrGlnThrLeuAsnThrLeuArgGlnAsnAlaGlnIleLeuValIleCys 106
QY 570 ----- 570
Db 107 GluLysLysCysValGluLeuLysPheThrLeuAspPheThrGlnAsnSerGluLysAsn 126
QY 571 -----TTGGAA 576
Db 127 PheIleIlePhePheArgLysPheMetSerPheArgLeuLysLysLeuArgPheLeuLys 146
QY 577 GTCTTTGGTCCATTA----- 591
Db 147 SerPheThrProIlePheLeuThrPheIleThrGlnAsnValThrLysIleAlaIleIle 166
QY 591 ----- 591
Db 167 SerGlnPhePheHisThrLysAsnAlaThrLysSerHisArgThrHisPheSerValIle 186
QY 592 -----CCAGCTTTTGTACGA----- 606
Db 187 ThrLysAsnAlaLysAsnProAspPheProLysAsnIlePheArgArgProLeuAlaSer 206
QY 607 -----GTCTGGGTGGAACCTCCTCATATAATACGCCCTCGAGCTACTGCTGTGATATCCCTG 660
Db 207 SerPheMetTrpValSerThrCysLeuSerTyrProAlaValLeuAlaIleGlnAlaIle 226
QY 688 -----TTTTTTTCAATGTGAAATCCCTGAACTTGGATCAAGCTC 729
Db 247 PheSerLysPheLeuAsnPheLeuPheLysTyrThrValSerTyrPheIleLeuLysIle 266
QY 730 -----ATTACA----- 735
Db 267 SerPhePheLysHisLeuIleLysPhePheGlnGlyLeuAspSerTrpIleThrIle 286
QY 736 -----GCTGTGGGCATAACTGTA-----GTG 756
Db 287 AspGluAsnTrpArgPheMetThrTyrArgLeuValGlyPheSerMetLeuTrpProLeu 306
QY 757 ATGTCCTCTAAATAGCATGATGTCAGTGTGAGTGGAGCGCC---CGGATCCAGATTTTCTTAACC 813
Db 307 MetLeuLeuAsnPhePheSerLeuLysLysValAlaGlyAlaPheGlnIleValAlaThr 326
QY 814 TTTTGCAGGCTCACAGCAATCTGTGATAATTATAGTCCCTGGAGTTATGCAGCTAATT--- 870
Db 327 AlaIleLysLeuIleValAlaSerIleIleIleThrGlyLeuTyrHisIleIlePhe 346
QY 871 AAAGGTCAAACGCAGAACCTTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACGCGG 930
Db 347 LysGlnGlnThrGlnAsnPheLysAsnSerPheThrGlySerAspTrpAsnProGlyAsn 366

QY 931 TTGCCACTGGCTTTTATTATGGAATGATGATATGATATGCTGGCTGGTTTACCTCAACTTT 990
 Db 367 LeuValLeuGlyValTyrSerGlyLeuPheAlaTyrAsnGlyTrpAspValLeuAsnPhe 386
 QY 991 GTTACTGAAGAAGTAGAAAACCCCTGAAAAACCATTCCTCCCTTGCAATATGTATATCCATG 1050
 Db 387 GlyAlaGluGluIleGluAsnProArgArgThrLeuProIleAlaAlaIleSerGlyIle 406
 QY 1051 GCCATTGTCACCATTTGGCTATGTCTGACAAAATGGCGCTACTTTACGACCATTAATGCT 1110
 Db 407 AlaIleSerAlaThrValPheIleLeuMetAsnValSerTyrPheSerValLeuSerVal 426
 QY 1111 GAGGAGCTGCTGCTTTCAATGCGAGTGGCAGTG-----ACCTTTTCT 1152
 Db 427 GluAspPheLysAsnSerProAlaValAlaValArgThrPhePheIleThrPheAla 446
 QY 1153 GAGCGGCTACTGGGAAATTTCTCATATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTT 1212
 Db 447 GluArgThrLeuGlyAspPheHisTyrAlaIleProPheLeuIleSerLeuLeuIle 466
 QY 1213 GGCTCCATGAACGGTGGTGTGTTTGTCTCTCCAGTTATTCTATGTTGCGTCTCGAGAG 1272
 Db 467 GlySerMetAsnThrThrIlePheAlaCysSerArgTyrMetTyrSerGlyAlaGlnGln 486
 QY 1273 GGTCACTTCCAGAAATCCTCTCCATGATTGATGTCGCGCAAGCACACCTCCTCTACCAGCT 1332
 Db 487 SerValMetProThrProLeuArgGlyIleHisArgThrArgSerProArgLeuAla 506
 QY 1333 GTTATTGTTTGCACCCCTTGACAATGATAATGCTCTCTCTCTGGAGACCTCGACAGTCTT 1392
 Db 507 ValPheAlaGluIleLeuIleAlaIleCysLeuSerPheIleGlyAsnLeuAspGlnLeu 526
 QY 1393 TTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATT 1452
 Db 527 IleSerTyrMetSerPheAlaLeuTrpSerGlnArgThrCysThrGlnValGlyPheIle 546
 QY 1453 TATCTTCGA-----TACAAATGCCAGATATGCATCGTCTCTTCAAGGTGCCA 1500
 Db 547 TyrPheLysLeuArgGlyThrLeuLysThrGlnAsp-----SerPheGlnValPro 563
 QY 1501 CTGTTTATCCAGCTTTGTTTCTTCCATGCTCTCTCTCATGTTGCCCTTTCCCTCTAT 1560
 Db 564 IlePheValProValValPheLeuGlyIleCysIleAlaLeuLeuValIleProIleThr 583
 QY 1561 TCGGACCCATTTAGTACAGGATTGGCTTCGTCATCATCTGCTGAGTCCCTGCGTAT 1620
 Db 584 GlnAsnTyrHisValAlaIleTyrGlyValSerMetThrIleGlyGlyAlaIleIleTyr 603
 QY 1621 TATCTCTTTATATATGGGACAAAGAACCCAGGTGGTTTATAGATAATGTACAGAAAAATA 1680
 Db 604 LeuIlePheIlePheProAsnThrLeuProIlePheLeuHisLysIleAsnAsnSerIle 623
 QY 1681 ACCAGAACATTACAAATAATCTGGAAGTTGTA-----CCAGAAGAAGAT 1725
 Db 624 ValLysPheAlaGlnIleIlePheAsnCysValIleGluProTyrGluAsp 640

RESULT 7

T31554
 hypothetical protein Y53H1C.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T31554
 R:White, S.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21044
 A:Accession: T31554
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-541 <WIL>
 A:Cross-references: EMBL:AL117201; PIDN:CAB55063.1; CESP:Y53H1C.1
 A:Experimental source: clone Y53H1C
 C:Genetics:
 A:Gene: CESP:Y53H1C.1

A; Introns: 31/1; 90/3; 127/1; 154/1; 198/1; 249/2; 325/2; 371/2; 433/3; 467/1; 510/1
 Alignment Scores:
 Pred. No.: 4.35e-51 Length: 541
 Score: 665.50 Matches: 156
 Percent Similarity: 55.93% Conservative: 108
 Best Local Similarity: 33.05% Mismatches: 169
 Query Match: 16.15% Indels: 39
 DB: 2 Gaps: 9

US-09-667-170A-440 (1-2239) x T31554 (1-541)

QY 376 GTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCAATCTCTCTAAGGCGGTG 435
 Db 26 MetSerTyrValIleAlaAsnIleIleGlyAlaGlyIlePheIleThrProGlyThrVal 45
 QY 436 CTCCAGAACACGGGCGCAGCGTGGGCAATGCTCTGTGACCATCTGGACGGTGTGGGCTCCTG 495
 Db 46 LeuLysLeuAlaLeuThrAsnGlyMetAlaLeuValValTrpLeuGlyCysGlyLeuIle 65
 QY 496 TCACTATTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGT 555
 Db 66 SerLeuIleGlyGlyIleCystyrIleGluLeuGlyThrSerIleArgAspProGlyCys 85
 QY 556 CATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTG 615
 Db 86 AspPheAlaTyrAsnValTyrValGlyTrpGluGlyIleAlaPheSerPheMetTrpVal 105
 QY 616 GAACCTCCTCATATAACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGCTAC 675
 Db 106 GlyValIleMetSerPheProAlaSerAlaAlaValGlnAlaGlnThrPheGlyGlnTyr 125
 QY 676 ATTCCTG-----GAACCATTTTATTCAATGTGAA 705
 Db 126 IleValAlaGlyLeuSerProIleTrpHisLeuGlySerProTyrAspValIleLeuGlu 145
 QY 706 ATCCCTGAACCTTGGGATCAAGCTCATTTACAGCTGTGGGCAATACTGTAGTATGGTCTTA 765
 Db 146 -----ArgGlyLeuGlyPheAlaLeuIleIleLeu 156
 QY 766 AATAGCATGAGTGC-----AGCTGGAGCGCGGATCCAGATTCTTCTTAACC 813
 Db 157 ThrValLeuAsnLeuTyrAlaIleAspLysTyrAlaSerLysPheGlnIlePheValThr 176
 QY 814 TTTGCAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATT--- 870
 Db 177 IleAlaLysLeuLeuSerLeuAlaIleIleIleValThrGlyPheTrpTyrLeuIleVal 196
 QY 871 AAAGGTCAAACGCAGAACTTTAAAGACGCCCTTTTCA-----GGAAGAGATTCAAGT 921
 Db 197 LysGlyGluThrGluHisPheLysAspAlaPheThrProLeuProAsnGluLysTyrAsp 216
 QY 922 ATTACGGGTTGCCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTAC 981
 Db 217 IleGlyGlnIleSerLeuAlaPheTyrGlyAlaLeuTrpSerPheAlaGlyTrpAspIle 236
 QY 982 CTCAACTTTGTTACTGAAGAAGTAGAAAAACCCCTGAAAAAACCATTTCCCTTGCATATGT 1041
 Db 237 LeuAsnTyrGlyThrProGluIleLysAsnProArgArgThrMetProIleAlaLeuLeu 256
 QY 1042 ATATCCATGGCCATTTGTCACCATTTGGCTATGCTGACAAAATGTGGCCTACTTTACGACC 1101
 Db 257 GlyGlyIleSerAlaValThrLeuValTyrMetAlaMetAsnValSerTyrMetThrVal 276
 QY 1102 ATTAATGCTGAGGAGCTGCTGCTTTTCAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTA 1161
 Db 277 LeuAspThrGluThrLeuLysAsnSerSerAlaValAlaAlaAspPheAlaArgIleThr 296
 QY 1162 CTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCTGCTTGGCTCCATG 1221
 Db 297 LeuGlyAspPheSerTyrAlaIleProPheMetIleSerIleLeuLeuIleGlyThrLeu 316
 QY 1222 AACGGTGGTGTGTTGCTGTCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGGTCACTT 1281

Db 317 AsnSerAsnIlePheCysGlySerArgPheThrHisAlaAlaArgGluGlyHisLeu 336
QY 1282 CCAGAAATCCTCTCCATGATTTCATGTCGCAAGCACACTCTCTACCAGCTGTTATGTT 1341
Db 337 ProThrPheLeuSerCysIleAsnAlaGluSerAsnSerProArgAlaAlaLeuPhe 356
QY 1342 TTGCACCCCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTC 1401
Db 357 GlnLeuIleCysThrIleAlaValThrPheVal---AspThrGluSerLeuIleThrTyr 375
QY 1402 CTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATCTTCGA 1461
Db 376 ValThrPheValMetPheGlyGlnArgValPheThrMetAlaAlaLeuLeuTrpIleArg 395
QY 1462 TACAAATGCCAGATATGCAT--CGTCCTTTCAAGGTGCCACTGTTCATCCAGCTTTG 1518
Db 396 TyrArgAsnIleProValHisProAspAlaIleArgValProLeuIlePheSerIleLeu 415
QY 1519 TTTTCCTTCACATGCCTCTTCATGGTGGCCCTTTCCCTCTATTCCGACCCATTT----- 1572
Db 416 Phe-----PheMetIleThrIleAlaLeuValThrProPheIleGlu 430
QY 1573 -----AGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTAT 1623
Db 431 AspPheThrThrIleValGlyValGlyLeuValLeuMetGlyPheLeuLeuTyrMet 450
QY 1624 CTCTTTATTATATGGGACAAGAAACCCAGGTGGTGTAGAAATAATGTCAGAGAAATAACC 1683
Db 451 IlePheMetLysProLysGlnLeuProGlyPheLeuTyrArgPheAsnAspGlyIleThr 470
QY 1684 AGAACATTACAAATAATACTGGAAGTTGTACAGAA 1719
Db 471 ArgValThrCysLysIleLeuPheThrThrProAsp 482

RESULT 8

T16854
hypothetical protein T13A10.10 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T16854

R;Nelson, J.

submitted to the EMBL Data Library, April 1996

A;Description: The sequence of C. elegans cosmid T13A10.

A;Reference number: Z18590

A;Accession: T16854

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-562 <NEL>

A;Cross-references: EMBL:U56963; NID:g1293813; PID:g1293821; PIDN:AAB38125.1; GSPDB:GN00

A;Experimental source: strain Bristol N2; clone T13A10

C;Genetics:

A;Gene: CESP:T13A10.10

A;Map position: 4

A;Introns: 23/2; 107/3; 140/1; 167/1; 211/1; 259/2; 335/2; 381/2; 500/1; 540/1

Alignment Scores:

Pred. No.: 3.11e-47 Length: 562
Score: 622.50 Matches: 154
Percent Similarity: 54.03% Conservative: 101
Best Local Similarity: 32.63% Mismatches: 180
Query Match: 15.11% Indels: 37
DB: 2 Gaps: 9

US-09-667-170A-440 (1-2239) x T16854 (1-562)

QY 397 ATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCTCCAGAACACGGGACGGTG 456

Db 50 IleIleGlyAlaGlyIlePheIleThrProGlyProIleLeuGlnTyrThrPheSerAsn 69

QY 457 GGCATGTCTGTACCATCTGGACGGTGTGTGGGTCTCTCACTATTGGAGCTTTGTCT 516

Db 70 GlyLeuAlaLeuLeuValTrpIleGlyCysGlyLeuIleSerLeuIleGlyIleCys 89

QY 517 TATGCTGAATTCGGAACAACACTATAAAGAAATCTGGAGTCAATTACACATATATTTTGGAA 576
Db 90 TyrIleGluLeuGlyThrSerIleHisAspProGlyCysAspPheAlaTyrThrValTyr 109
QY 577 GTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATAATACGCCCT 636
Db 110 ValGlyTrpGluGlyIleAlaPheSerPheMetTrpValGlyValIleMetSerPhePro 129
QY 637 GCAGCTACTGCTGTGATATCCCTGGCATTT-----GGACGCTACATTTCTGGAACCATTT 690
Db 130 AlaSerAlaAlaValGlnAlaLeuThrPheValAlaGlyMetAlaProIleTrpProLeu 149
QY 691 TTTATTCAATGTGAAATCCCTGAACTTGGCATCAAGCTCATTTACAGCTGTGGGCATAACT 750
Db 150 -----GluHisProTrpAspGlyIle---IleGluLysGlyLeuGlyPheAla 164
QY 751 GTAGTGTGTCCTAAATAGCATGAGTGTCT-----AGCTGGAGCGCCCGGATC 798
Db 165 LeuIleIleValLeuThrIleLeuAsnLeuTyrAlaIleAspLysTyrAlaSerLysPhe 184
QY 799 CAGATTTTCTTAACCTTTTGCAGCTCACAGCAATTTGATAATTTAGTCCCTGGAGTT 858
Db 185 GlnIleValValThrIleAlaLysMetLeuSerLeuAlaIleIleIleValThrGlyPhe 204
QY 859 ATGCAGCTAATT---AAAGGTCAAACGCAGAACTTTAAAGACGCTTTTTCAGGAAGAGAT 915
Db 205 TyrTyrLeuIlePheLysGlyGlnThrGluTyrLeuGluHisProPheGluGlySerAsn 224
QY 916 TCAAGTATTACGCGGTGCCACTGGCTTTTATTTATGGAATGTATGATATGCTGGCTGG 975
Db 225 ThrAsnProGlyGlnIleSerLeuAlaPheTyrGlyAlaLeuTrpSerPheAlaGlyTrp 244
QY 976 TTTTACCTCAACTTTTGTACTGAAGAAGTAGAAACCCCTGAAAAACCATTTCCCTTGCA 1035
Db 245 AspIleLeuAsnPheGlyThrProGluIleArgAsnProArgTrpMetProIleAla 264
QY 1036 ATATGTATATCCATGGCATTGTCAACCATTTGGCTATGTGCTGACAAATGTGCCTACTTT 1095
Db 265 LeuLeuGlyGlyValSerIleValThrAlaValTyrLeuAlaMetAsnIleSerTyrMet 284
QY 1096 ACGACCATTAATGCTGAGAGCTGCTGCTTTTCAAATGTCAGTGGCAGTGACCTTTCTGAG 1155
Db 285 ThrValLeuThrProGlyGlnIleMetAsnSerThrAlaValAlaAlaAspPheAlaGln 304
QY 1156 CGGCTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTTGTTGCTCTCTCTGCTTTGGC 1215
Db 305 IleThrLeuGlyGlyPheSerTyrAlaIleProPheMetIleAlaLeuLeuIleGly 324
QY 1216 TCCATGAACGGTGGTGTGTGTGCTGTCTCCAGGTTATTTCTATGTTGCGTCTCAGAGGGT 1275
Db 325 ThrLeuAsnSerAsnIlePheCysGlySerArgPheThrHisAlaAlaAlaArgGluGly 344
QY 1276 CACCTTCCAGAAATCCTCTCCATGATTTCATGTCCGCAAGCACACTCTCTACAGCTGTT 1335
Db 345 HisLeuProThrPheLeuSerCysIleAsnGluGluSerAsnSerProArgAlaAlaLeu 364
QY 1336 ATTGTTTGCACCCCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTG 1395
Db 365 LeuPheGlnLeuValCysThrValValThrPheIle---AspThrAsnSerLeuIle 383
QY 1396 AATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTGTGGGCTGATTAT 1455
Db 384 AsnTyrValSerPheValMetPheGlyGlnArgValPheThrMetThrAlaLeuMetTrp 403
QY 1456 CTTTCGATACAAATGCCAGATATGCAT---CGTCTTTTCAAGGTGCCACTGTTTCATCCCA 1512
Db 404 IleArgTyrArgAsnIleProValHisProAspAlaIleArgValProLeuIlePheSer 423
QY 1513 GCTTTGTTTTCCTTCACATGCCCTCTTCATGGTTGCC----- 1548
Db 424 PheLeuPheTrpLeuIleThrIleAlaLeuValValProPheIleGluGluPheThr 443

QY 1549 -----CTTCCCTCTATTCCGGAC 1566
Db 444 ValGlyValThrAspLeuProAsnLeuGluLysPheGlnAsnValGlnLeuIleSerGlu 463
QY 1567 CCATTT---AGTACAGGATTGGCTTCGTCATCACTCTGAGTGGAGTCCCTGCTATTAT 1623
Db 464 LysPheGlnGlnThrIleValGlyValGlyLeuValLeuMetGlyValPheLeuTyrIle 483
QY 1624 CTCTTTATTATATGGACAAGAACCCAGGTGGTTTGAATAATGTACAGAGAAATAACC 1683
Db 484 IlePheMetLysProMetLysLeuProGluPheLeuIleArgPheAsnAspSerMetThr 503
QY 1684 AGAACATTACAAATAATACTGGAGTTGTACCAGAA 1719
Db 504 ArgIleThrCysLysIleLeuPheThrThrProAsp 515

RESULT 9
T24837
hypothetical protein T11F9.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24837
R:Lennard, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19941
A:Accession: T24837
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-440 <WIL>
A:Cross-references: EMBL:Z74042; PIDN:CAA98529.1; GSPDB:GN00023; CESP:T11F9.4
A:Experimental source: clone T11F9
C:Genetics:
A:Gene: CESP:T11F9.4
A:Map position: 5
A:Introns: 41/3; 237/2; 269/3; 386/1

Alignment Scores:
Pred. No.: 5.18e-42 Length: 440
Score: 564.00 Matches: 136
Percent Similarity: 48.60% Conservative: 90
Best Local Similarity: 29.25% Mismatches: 159
Query Match: 13.69% Indels: 80
DB: 2 Gaps: 10

US-09-667-170A-440 (1-2239) x T24837 (1-440)

QY 352 AGGAAAGTCACCTTTACTGAGGGGAGTCCATATATCATTTGGACCATCATTTGGAGCAGGA 411
Db 8 GlnLysMetGlyLeuLeuGlyAlaIleSerTyrIleValGlyAsnIleValGlySerGly 27
QY 412 ATCTTCATCTCTCTAAGGGCGTGTCCAGAACACGGGAGCGTGGCATGTCTCTGACC 471
Db 28 IlePheIleThrProThrSerIleIleGluAsnValAsnSerIle----- 42
QY 472 ATCTGGACGGTGTGGGGTCTGTCTCACTATTGGAGCTTTGTCTTATGCTGAATGGGA 531
Db 43 -----GlnIleGlnGlyAla-----Gly 48
QY 532 ACAACTATAAGAAATCTGGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTA 591
Db 49 ValGluLeuAspGluThrSer----- 55
QY 592 CCAGCTTTGTACGAGTCTGGTGGAACTCCTCATATFACGCCCTGCAGCTACTGCTGTG 651
Db 56 -----GluPheTrpAlaLysLysLeuLeu-----GlyPheSerLeuIle 68
QY 652 ATATCCCTGGCATTTGGACGGCTACATCTCGGAACCATTTTATTCAATGTGAAATCCCT 711
Db 69 IleLeuLeuMetPheMetAsnPhePheSerLeuLysThrPheValGln----- 84
QY 712 GAACCTGGCATCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTATGGTCTCTAAATAGC 771
Db 84 ----- 84

QY 772 ATGAGTGTCTAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTTCGAAGCTCACAGCA 831
Db 85 -----ArgPheSerIleLeuAlaSerLeuAlaLysIleAlaAla 97
QY 832 ATTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATT---AAAGTCAACGCAGAAC 888
Db 98 ThrLeuLeuIleIleIleThrGlyPheTyrTyrLeuIlePheLysHisTrpLysGlnAsn 117
QY 889 TTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACGGCGTTGCCACTGGCTTTTAT 948
Db 118 LeuGluGluProPheLysGlySerAsnTrpAsnProGlyProPheValAsnAlaLeuPhe 137
QY 949 TATGGAATGTATGCATATGCTGGCTGGTGGTTTACCTCAACTTTTGTACTGAAGAAGTAGAA 1008
Db 138 AlaGlyLeuPheSerTyrAspGlyTrpAspIleLeuAsnPheGlyAlaGluIleGlu 157
QY 1009 AACCTTGAAAAACCATTTCCCTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1068
Db 158 AsnProLysArgThrMetProLeuSerIleIleIleGlyMetThrCysIleGlyValIle 177
QY 1069 TATGTCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCA 1128
Db 178 TyrValAlaValAsnValAlaTyrSerIleValLeuSerProThrGluMetIleAlaSer 197
QY 1129 AATGCAGTGGCAGTACCTTTTCTGACGGCTACTGGGAAATTTCTCATAGCAGTTCGG 1188
Db 198 AsnAlaValAlaIleAspPheAlaAsnLysThrLeuGlyAlaAlaAlaPheValPro 217
QY 1189 ATCTTTGTTGGCCTCTCCTGCTTGGCTCCATGAACGGTGGTGGTGGTGGTGGTGGTGG 1248
Db 218 ValMetValAlaIleLeuLeuIleGlySerLeuAsnSerThrMetPheSerAlaSerArg 237
QY 1249 TTATTCTATGTGCGTCTCGAGAGGTCACTTCCAGAAATCCTCTCCATGATTTCATGTC 1308
Db 238 TyrLeuGlnAlaValSerArgGlnGlyHisIleProSerAlaIleSerGlyIleAlaPro 257
QY 1309 CGCAAGCACACTCTCTACCGCTGTTATTGTTTGGACCCCTTTGACAATGATAATGCTC 1368
Db 258 AsnCysAspSerProArgValAlaLeuLeuValHisIleLeuIleAlaIleAlaValSer 277
QY 1369 TTCTCTGGAGACCTCGACAGTCTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTATT 1428
Db 278 PheLeuGlyAspProAspLysLeuIleAsnTyrValAlaPheAlaGlnTrpSerGlnArg 297
QY 1429 GGGCTGGCAGTGTGGGCTGATTTATCTTCTGATACAAATGCCAGATATGCATCGTCTC 1488
Db 298 AlaPheThrMetSerAlaLeuLeuTyrLeuArgIleArgGlyArgProArgHisProAsp 317
QY 1489 ---TTCAAGGTGCCACTGTTTCATCCCGAGCTTTGTTTCCCTTCACATGCCCTTTCATGG 1545
Db 318 ArgIleGlnLeuProIleIleMetProIleLeuPhePheLeuValCysThrSerMetVal 337
QY 1546 GCCCTTCCCTCTATTTCGGACCCCATTTAGTACAGGGATTGGCTTCGTATCATCTCTGACT 1605
Db 338 ValIleSerIleIleAspAspPheLysSerSerAlaValGlyLeuGlyIleLeuLeuGly 357
QY 1606 GGAGTCCCTGCGTATTATCTCTTTTATTATATATGGAGCAAG-----AAACCCAGGTGG 1656
Db 358 GlyLeuIleIlePheIleIlePhe---ValTrpAspArgAlaLeuProSerSerHisThr 376
QY 1657 TTAGA-----ATAATGTACAGAGAAATAACAGAACATTACAAATAATACTG 1704
Db 377 PheArgAsnAlaThrHisValIleAsnGluGluSerThrLysPheMetGlnIleIlePhe 396
QY 1705 GAAGTTGTACCAGAA 1719
Db 397 AsnValValProGlu 401

RESULT 10
B69855
amino acid permease homolog ykba - Bacillus subtilis
C:Species: Bacillus subtilis

[illegible]

QY 484 TGTGGGGTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATGGGAACAACATAAAG 543
Db 67 GlyGlyIleValSerIleIleGlyAlaLeuGlyAlaGluLeuAlaThrThrTyrPro 86
QY 544 AAATCTGGAGGTCAATTACACATATATTTTGAAGTCTTTTGGTCCATTACCAGCTTTTGT 603
Db 87 AsnValGlyGlyAlaTyrTyrTyrLeuLysArgAlaPheGlyGlnAsnThrAlaPheLeu 106
QY 604 CGAGTCTGGGTGAACCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCA 663
Db 107 PheAlaTyrAlaArgLeuThrValIleGlnThrGlySerIleAlaLeuAlaAlaPheVal 126
QY 664 TTTGGAGCTACATCTCTGAACCATTTTTCATCAATGTGAATCCCTGAACCTTGCATC 723
Db 127 PheGlyAspTyrAlaSerGluIleTyrPheSerSer 142
QY 724 AAGCTCATACAGCTGTGGGATAACTGTAGTGATGCTCTAAATAGCATGAGTGTGAGC 783
Db 143 SerMetTyrAlaAlaValIleIleAlaLeuThrIleLeuAsnIleLeuGlyLeuHis 162
QY 784 TGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATCTCTGATAAT 843
Db 163 GlnGlyLysTyrThrGlnAsnLeuThrAlaAlaGlnValLeuGlyLeuLeuVal 182
QY 844 ATAGTCCCTGGAGTATGCGAGCTAATTAAGGTCAACGCGAGAACTTTAAAGACGCTTT 903
Db 183 ValLeuPheGly-----LeuAlaSerThrAlaAsnSerAlaAsnSerAlaVal 198
QY 904 TCAGGAAGAGATTCAGTATTACGCGGTGGCCACTGGCTTTTATTTATGGAATGTATGCA 963
Db 199 SerProGluProSerSerSerGlySerTyrGlyLeuAlaMetValPheValLeuLeuSer 218
QY 964 TATGCTGGTGGTGTATACCTCAACTTTGTTACTGAAGAGTAGAAACCCCTGAAAAACC 1023
Db 219 TyrGlyGlyTyrAsnGluAlaAlaTyrIleSerAlaGluIleLysAsnArgGlnArgAsn 238
QY 1024 ATTCCTCTTGCATATGATATATCCATGCGCATTTGTCACCATTTGGCTATGCTGACAAAT 1083
Db 239 IleLeuArgSerLeuMetTyrSerIleGlyIleIleThrAlaIleTyrLeuLeuIleAsn 258
QY 1084 GTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAATGCGAGTGGCAGTG 1143
Db 259 LeuAlaPheLeuArgGlyLeuGlyLeuAlaAsnMetAlaAsnSerSerAlaValAlaAla 278
QY 1144 ACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCA---GTTCCGATCTTTGTGCC 1200
Db 279 AspLeuMetArgAlaValTyrGlyTyrProGlyValValPheIleSerLeuLeuIleAla 298
QY 1201 CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTT 1260
Db 299 IleCysAlaLeuGlyThrIleAsnAlaSerIlePheThrGlyAlaArgThrAsnTyrAla 318
QY 1261 GCGTCTCGAGAGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGCAAGCACACT 1320
Db 319 LeuGlyGlnAspPheAsnLeuPheGlyPheMetGlySerTyrArgGlnIleProSerThr 338
QY 1321 CCTCTACAGCTGTTATTTGTCACCCCTTTGACAATGATAATGCTCTTCTCTGGA--- 1377
Db 339 ProAlaThrAlaLeuLeuValGlnGlyAlaIleAlaLeuAlaLeuValValLeuGlyThr 358
QY 1378 -----GACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTT 1425
Db 359 PheThrArgLysGlyPheGluThrMetValAspTyrThrAlaProValPheTyrPhePhe 378
QY 1426 ATGGGCTGGCAGTGTGCTGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGT 1485
Db 379 PheLeuLeuSerGlyIleSerLeuLeuIleLeuArgGlnLysGluProHisIleProArg 398
QY 1486 CCTTTCAAGGTGCCACTGTTT-----ATCCCAGCTTTGTTTCTTCACATGCTCTTTC 1539
Db 399 ProPheArgValProPheTyrProIleThrProLeuLeuPheCysAlaValCysGlyTyr 418
QY 1540 ATGGTTGCCCTTCCCTCTATTCGGACCCATTTTAGTACAGGGATTGGCTTCGTC----- 1593

Db 419 Leu-----LeuTyrSerSerValValTyrThrAsnValGlyAlaValValGly 434
QY 1594 -----ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGGACAAG 1644
Db 435 ValLeuValValIleAlaGlyValPro-----LeuLeuPheTyrAsnArg 449
RESULT 13
B69172
cationic amino acid transporter related protein - Methanobacterium thermoautotrophicum
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Aug-1999
C;Accession: B69172
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني,
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: B69172
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-424 <MTH>
A;Cross-references: GB:AE000837; GB:AE000666; NID:g2621613; PIDN:AAB85052.1; PID:g262
C;Genetics:
A;Gene: MTH546
C;Superfamily: arginine permease
Alignment Scores:
pred. No.: 9.92e-28 Length: 424
Score: 404.50 Matches: 119
Percent Similarity: 45.91% Conservative: 83
Best Local Similarity: 27.05% Mismatches: 205
Query Match: 9.82% Indels: 33
DB: 2 Gaps: 13
US-09-667-170A-440 (1-2239) x B69172 (1-424)
QY 337 AAAGTGCAGCTGAAGAGAGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGCACC 396
Db 2 LysProAsnLeuArgGluLeuGlyLeuPheAspAlaValAsnLeuValGlyThr 21
QY 397 ATCATTTGGAGCAGGAATCTTCATCTCTCTAAAGGCGTGTCTCCAGAACACGGGCGCGTG 456
Db 22 IleValGlyAlaAspIleTyrIleVal-----AlaAlaTyrGlyAlaGlySerLeu 38
QY 457 GGC---ATGCTCTCTGACCATCTGGAGGTGTGTGGGCTCCTGTCTCATTGTTGGAGCTTTG 513
Db 39 GlyProAlaSerIleLeuAlaTyrLeuLeuAlaGlyLeuMetAlaLeuIleAlaLeu 58
QY 514 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCTTACACATATATTTTG 573
Db 59 ValPheSerGluAlaSerAlaMetLeuProArgThrGlyGlyProTyrValTyrAlaGly 78
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAATACGC 633
Db 79 GluAlaLeuGlyArgPheThrGlyPheIleThrGlyTrp---SerLeuTyrValSerSer 97
QY 634 CCTGCAGTACTGCTGTGATATCCCTGGCATTGTTGGAGCTACATTCTTGGAAACCATTTT 693
Db 98 TrpValAlaIleAlaValPheProLeuAlaPheIleTyrTyrLeu-----GluTyrPhe 115
QY 694 ATTCATATGTAATCCCT---GAAGTTCGATCAAGCTCATACAGCTGTGGGCATAACT 750
Db 116 IleProLeuAspProProAlaGluAlaValIleLysValLeuPheIleLeuSerLeuThr 135
QY 751 GTAGTGTGCTCCTAAATAGCATGATGATGTCAGCTGGAGCGCCCGGATCCAGATTTCTTA 810
Db 136 IleIle-----AsnIleAlaGlyValGlyArgAlaGlyLysValAsnAspIleLeu 152
QY 811 ACCTTTTGCAGCTCACAGCAATTTCTGATAAATATATAGTCCCTGGAGTTATGCAGCTAAT 870

Db 153 ThrIleLeuLysValAlaProValLeuLeuPheAlaValLeuGlyAlaIleHisLeuAla 172
QY 871 AAAGGTCAA-----ACGAGAACTTTAAAGACGCCTTTTCA-----GGAAGA 912
Db 173 LeuAsnProGlyLeuLeuValSerAsnTyrThrProAlaAlaProMetGlyLeuGlyAla 192
QY 913 GATTCAAGTATTACGGGTTGGCCACTGGCTTTTATATGAATGTATGCATATGCTGGC 972
Db 193 LeuGlyThrValThrValLeuVal-----PheTrpAlaTyrValGly 206
QY 973 TGGTTTACCTCAACTTTGTACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTCCTT 1032
Db 207 PheGluLeuValThrValProAlaAspGluValArgAspProGluArgThrIleProLeu 226
QY 1033 GCAATATGTATATCCATGGCCATTGTCAACCATTTGGCTATGTCTGACAAATGTGGCCTAC 1092
Db 227 SerIleThrLeuGlyMetIlePheValThrLeuPheTyrIleLeuThrAsnAlaValIle 246
QY 1093 TTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCT 1152
Db 247 LeuGlyLeuValProTrpArgValLeuAlaSerSerThrAlaProLeuThrValAlaGly 266
QY 1153 GAGCGGCTACTGGGAATTC---TCATTAGCAGTTCGGATCTTTGTTGCCCTCTCCTGC 1209
Db 267 TyrSerLeuMetGlyGlyIleGlyAlaLeuLeuThrAlaGlyAlaValPheSerIle 286
QY 1210 TTTGGCTCCATGAACGGTGGTGTGTTTGTCTCTCCAGGTTATCTATGTTGGCTCTCGA 1269
Db 287 AlaGlySerGluAlaGlyMetLeuThrThrAlaArgLeuLeuPheAlaMetSerGlu 306
QY 1270 GAGGCTCACCTTCAGAAATCCTCTCCATGATTCATGTCCGGAAGCACACTCCTCTACCA 1329
Db 307 AspGlyPheLeuProGlyPheLeuSerArgValHisArgPheGlyThrProHisMet 326
QY 1330 GCTGTATTGTTTGCACCCCTTTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGT 1389
Db 327 SerIleLeuValGlnAsnLeuThrAlaLeuLeuAlaLeuThrGlyThrValSerGly 346
QY 1390 CTTTGAATTTCTCAGTTTGGCCAGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTG 1449
Db 347 LeuIleGluLeuSerValValThrLeuLeuLeuProTyrAlaValThrCysIleSerLeu 366
QY 1450 ATTTATCTTCGATACAAATGCCAGATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATC 1509
Db 367 AlaIleLeuArgArgArg---AspGlySerGlyIleProLeuLysSerValLeuGlyVal 385
QY 1510 CCAGCTTTGTTTCTTCCATATGCTTTCATGTTGGCCTTTCCCTCTATTCGGACCCA 1569
Db 386 -----LeuValCysIleTyrLeuLeu-----MetAsnThrThrPro 397
QY 1570 TTTAGTACAGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATATCTCTTT 1629
Db 398 SerThrThrAlaTrpGlyLeuLeuLeuIleLeuSerGlyAlaProLeuTyrLeuIlePhe 417
RESULT 14
S61943
methionine transport protein, high affinity - yeast (Saccharomyces cerevisiae)
N:Alternate names: methionine permease; protein G4340; protein YGR055w
C:Species: Saccharomyces cerevisiae
C>Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000
C:Accession: S61943; S64349
R:Isnard, A.D.; Thomas, D.; Surdin-Kerjan, Y.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61943
A:Accession: S61943
A:Molecule type: DNA
A:Residues: 1-574 <ISN>
A:Cross-references: EMBL:U40316; NID:g1101906; PIDN:AAB63529.1; PID:g1101907
A:Experimental source: strain X2180-1A
R:Entian, K.D.; Rose, M.; Koetter, P.; Roehmer, A.; Sehrsam, I.; Hempel, S.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64335
A:Accession: S64349

A:Molecule type: DNA
A:Residues: 1-574 <ENT>
A:Cross-references: EMBL:272840; NID:g1323066; PIDN:CAA97055.1; PID:g1323067; MIPS:YG
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:MUP1
A:Cross-references: SGD:S0003287; MIPS:YGR055w
A:Map position: 7R
C:Superfamily: hypothetical protein YHL036w
C:Keywords: amino acid transport; transmembrane protein
F;92-108/Domain: transmembrane #status predicted <TM1>
F;140-156/Domain: transmembrane #status predicted <TM2>
F;183-199/Domain: transmembrane #status predicted <TM3>
F;208-224/Domain: transmembrane #status predicted <TM4>
F;296-312/Domain: transmembrane #status predicted <TM5>
F;318-334/Domain: transmembrane #status predicted <TM6>
F;343-359/Domain: transmembrane #status predicted <TM7>
F;390-406/Domain: transmembrane #status predicted <TM8>
F;457-473/Domain: transmembrane #status predicted <TM9>
F;494-510/Domain: transmembrane #status predicted <TM10>
Alignment Scores:
Pred. No.: 1.75e-24 Length: 574
Score: 368.50 Matches: 135
Percent Similarity: 45.61% Conservative: 104
Best Local Similarity: 25.76% Mismatches: 199
Query Match: 8.94% Indels: 86
DB: 2 Gaps: 22
US-09-667-170A-440 (1-2239) x S61943 (1-574)
QY 280 GACGCCCTTTTCAGGAAGAGACGCCTTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGAGAAA 339
Db 36 AspAlaAspAsnGlyAlaSerAspPheGluAlaGlyGlnGlnPheAla---ThrGluLeu 54
QY 340 GTGCAGCTGAAGAGAGAAAGTCACTTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATC 399
Db 55 AspGlnGlyGluLysGlnLeuGlyIleLeuSerCysIleGlyLeuIleCysAsnArgMet 74
QY 400 ATTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTCTCCAGAACACGCGCAGCTGGGC 459
Db 75 LeuGlyThrGlyValPheAlaValSerSerThrIleTyrThrLeuCysGlySerValGly 94
QY 460 ATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTCTCACTATTTGGAGCTTTGTCTTAT 519
Db 95 LeuAlaLeuIleMetTrpAlaValGlyAlaIleIleAlaIleSerGlyLeuTyrValTyr 114
QY 520 GCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCACTTACACATATATTTTGGAAAGTC 579
Db 115 MetGluPheGlyThrAlaIleProLysAsnGlyGlyGluLysAsnTyrLeuGluAlaIle 134
QY 580 TTTGGTCCATTACCACTTTTGTACGAGTCTGGGTGGAACTCTCATATAATACGCCCTGCA 639
Db 135 PheArgLys---ProLysPhe----- 140
QY 640 GCTACTGTGTGATATCCCTGGCATTTTGGACGCTACATTTCTGGAACCATTTTATTCAA 699
Db 141 -----PheIleThrCysMetTyrAlaAlaTyrIle-----PhePheLeuGly 154
QY 700 TGTGAAATCCCT-----GAACCTTGCATCAAGCTCATTTACAGCT----- 738
Db 155 TrpAlaAlaGlyAsnSerIleAsnThrAlaIleMetPheLeuThrAlaAlaAspThrGlu 174
QY 739 -----GTGGGCATAACTGTAGTG-----ATGCTCCTA 765
Db 175 ValThrLysTrpAsnGlnArgGlyIleGlyValAlaValPhePheAlaPheLeuIle 194
QY 766 AATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTC 825
Db 195 AsnSerLeuAsnValLysIleGlyLeuTyrLeuGlnAsnIleLeuGlyIlePheLysIle 214
QY 826 ACAGCAATTCTGATAATTATAGTCCCTGGAGTTATGCACGTAATTAAGGT----- 876

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87396
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-546 <STO>
A;Cross-references: GB:AE005673; NID:g13422508; PIDN:AAK23168.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC1184

Alignment Scores:
Pred. No.: 5.77e-23 Length: 546
Score: 351.50 Matches: 132
Percent Similarity: 39.96% Conservative: 89
Best Local Similarity: 23.87% Mismatches: 204
Query Match: 8.53% Indels: 128
DB: 2 Gaps: 18

US-09-667-170A-440 (1-2239) x D87396 (1-546)
QY 273 TGGCAGGAGCGCCTTTTCAGGAAGACAGACGCCTTTTCAGGAAGACAGACGCCTTTTCAGGAAG 332
Db 11 TrpGlyAsnAlaHisMetAlaGlyAsnArgLeuPheLeuLys-SerIleAlaSerIl 30
QY 333 AGAGAAA-----GTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTC 380
Db 30 eGlnLysGluAlaAlaHisSerGlnLeuLysArgThrLeuGlyProIleAsnLeuMetSe 50
QY 381 CATTATCATTTGGCACCACCATTTGGAGCAGGAATCTTCACTCTCTCTAAGGGCGTGTCCA 440
Db 50 rLeuGlyValGlyAlaIleIleGlyAlaGlyIlePheValLeuThrGlyGlnValAlase 70
QY 441 GAACACGGGCGAGCGTGGGCATGTCTCTGACCATCTGGACGGTGTGGGGTCTGTCACT 500
Db 70 rAlaAsnAlaGlyProAlaIleMetLeuSer---PheIleValAlaGlyIleAlaCysAl 89
QY 501 ATTTGGAGCTTTGTCTTATGTGTCATTTGGGAACAACTATAAGAAATCTGGAGGTCACTTA 560
Db 89 aLeuAlaGlyLeuCysTy rAlaGluLeuAlaSerThrMetProValSerGlySerAlaTy 109
QY 561 CACATATATTTTGGAAAGTCTTTGGTCCATTACACAGCTTTTGTACGAGTCTGGTGGAACT 620
Db 109 rThrTy rAlaTy rGlyThrLeuGlyGluValPheAlaTrpIleMetGlyTrpLeu---Le 128
QY 621 CCTCATATAACGCCCTCGAGCTACTGTCTGTATATCCCTGGCATTTGGACGCTACATCTT 680
Db 128 uValLeuGluTy rGlyValAlaAlaSerThrValAlaValGlyTrpSerGlyTy rValVa 148
QY 681 GGAA-----TCCCTCTATTCCGAC-----CCATTT----- 684
Db 148 iSerThrLeuHisAlaLeuGlyIleAsnPheProMetIleGlnValAlaGlyAlaAspAl 168
QY 685 -CCATTTTATTCAATGTGAATCCCTGAACCTTGG----- 720
Db 168 aProMetTrpAlaThrProLeuIleGlnAlaValAlaAlaProGlyGlyThrMetPh 188
QY 721 -----ATCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGTGCTCCT 764
Db 188 eAlaMetThrGlyThrLeuAsnLeuValAlaAlaIleGlyIleAlaMetValSerAlaLe 208
QY 765 AAATAGCATGAGTGTACAGCTGGAGCGCCCGGATCCAGATTTCTTAACCTTTTGGCAAGCT 824
Db 208 uLeuValValGlyValSerGluSerAlaAsnValAsnAlaIleValValIleLysVa 228
QY 825 CACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACACGCA 884
Db 228 lIleValLeuValThrPheIleAlaValGlyAlaGlnTy rIleAsnProAlaAsnTrpHi 248
QY 885 GAACCTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTACGCGGTTGCCACTGGCTTT 944
Db 885 GAACCTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTACGCGGTTGCCACTGGCTTT 944

Db 215 GlyIleValLeuPheIleSerIleThrGlyTrpValAlaLeuGlyGlyLeuLysAsp 234
QY 877 -----CAAACGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACG--- 927
Db 235 GlyTy rGlnSerHisAsnPheArgAsnAlaPheGluGlyThrGluThrAlaThrAlaTy r 254
QY 928 CGGTGCGCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTGTACCTCAAC 987
Db 255 GlyIleValAsnAlaLeuTy rSerValIleTrpSerPheValGlyTy rSerAsnValAsn 274
QY 988 TTTGTTACTGAAGAAGTAGAAAACCCCTGAAAAACCATTCCTTTCGCAATATGTATATCC 1047
Db 275 TyrAlaLeuGlyGluValLysAsnProValArgThrLeuLysIleAlaGlyProThrSer 294
QY 1048 ATGGCCATTTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCCTACTTTACGACCATTAA 1107
Db 295 MetValPheLeuAlaIleIleTy rIlePheValAsnIleAlaTy rPheAlaValValPro 314
QY 1108 GCTGAGGAGCTGCTCTTCAAAATGCAGTG---GCAGTGACCTTTTCTGAGCGGCTACTG 1164
Db 315 LysAspLysLeuIleSerSerLysLeuIleLeuAlaAlaAspPhePheAspIleValPhe 334
QY 1165 ---GGAAATTTCTATTAGCAGTCCGATCTTTGTTGCCCTCTCTCTGCTTTGGCTCCCATG 1221
Db 335 GlyGlyGlnAlaLysArgAlaAlaAlaLeuValGlyLeuSerAlaLeuGlyAsnVal 354
QY 1222 AACGGTGGTGTGTTGCTGCTCCAGGTTATTCTATGTTGGTCTCGAGAGGGTCACTT 1281
Db 355 LeuSerValIlePheSerGlnGlyArgIleIleGlnGlnLeuGlyArgGluGlyValLeu 374
QY 1282 CCA-----GAAATCTCTCCATGATTTCATGTCGCGCAAGCACACTCTCTA 1326
Db 375 PropheSerAsnPhePheAla-SerSerLysPropheAsnSerProMetValGlyLeuPh 394
QY 1327 CCAGCTGTTATGTTTGGACCCCTTTTGACAATGATAATGCTCTCTCTGGAGACCTCGAC 1386
Db 394 eGlnHisPheIleValCysThr--valThrIleLeuAlaProProGlyAspAlaTy r 413
QY 1387 AGTCTTTTG---AATTTCTCAGTTTTCAGGTTGGCTTTTATTGGGCTGGCAGTTGCT 1443
Db 414 LeuLeuValGlnAsnLeuIleSerTy rProMetAsnIleIleAsnPheAlaIleSerAla 433
QY 1444 GGGCTGATTTATCTT-----CGATACAAATGCCAGATATGCATCGTCTCTTC 1491
Db 434 GlyLeuLeuTrpIleTy rTrpGlnArgArgGlnGlyLysIleGluTrpAsnProIle 453
QY 1492 AAGGTGCCACTGTTTCATCCAGCTTTGTTTCCTTCACATGCCTCTTCATGGTTGCCCTT 1551
Db 454 LysAlaGlyValPheValThrGlyPhePheThrLeuSerAsnLeuTy rLeuIleIleAla 473
QY 1552 -----TCCCTCTATTCCGAC-----CCATTT----- 1572
Db 474 ProTy rValProProSerAsnGlyGluSerValTy rSerSerMetProTy rTrpIleHis 493
QY 1573 -----AGTACAGGATTTGGCTTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTAT 1623
Db 494 CysValIleAlaTrpGlyIlePhePhe-----PheGlyGlyVal-----Ty rTy r 508
QY 1624 CTCCTTATTATATGGGACAAG---AAACCCAGGTGG-----TTTAGAATAATGTACAGAG 1674
Db 509 -----ValValTrpAlaGlnLeuLeuProArgTrpGlyHisTy rLysLeuValSerLys 526
QY 1675 AAAATA 1680
Db 527 AspVal 528

RESULT 15
D87396
amino acid permease family protein CC1184 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: D87396
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.


```

Db 248 sPropheilleProGluProThrGlyGln-----ProGlyGluPh 261
QY 945 TTATTATGGA-----ATGTATGCATATGCTGGCTGGTT 977
Db 261 eGlyIleGlyGlyIlePheArgGlyAlaAlaIleIlePhePheAlaTyrValGlyPheG1 281
QY 978 TTACCTCAACTTGTACTGAAGAAGTAGAAAACCCCTGAAAAACCATTCCTTGGCAAT 1037
Db 281 uAlaValSerThrAlaAlaAlaGluAlaLysAsnProSerArgAspValProIleGlyI1 301
QY 1038 ATGTATATCCATGGCCATGTGCACCATTTGGCTAT-----GTGCTGACAAA 1082
Db 301 eLeuGlyAlaLeuIleIleCysThrLeuIleIleTyrMetAlaValAlaAlaValMetThrG1 321
QY 1083 TGTGGCCTACTTTACGACCATTAAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCAGT 1142
Db 321 yValValProphe-----ArgGluLeuAlaSerProAlaProIleAlaVa 336
QY 1143 GACCTTTCTGAGCGGCTA-----1161
Db 336 lAlaIle---AspArgMetGlyLeuGluTrpAlaAspIleProTyrAlaAlaAlaGluG1 355
QY 1162 -----CTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCC-- 1206
Db 355 yGlyLysLeuAsnLeuLeuSerPheAlaIleLysIleGlyAlaIleThrGlyLeuSerSe 375
QY 1207 -----TGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGCTCCAGGTT 1250
Db 375 rValMetLeuValLeuCysTyrGlyGln-----ThrArgI1 387
QY 1251 ATTCTATGTTGGCTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCCG 1310
Db 387 ePheTyrThrMetAlaArgAspGlyLeuLeuProLysValPheAlaGluIleHisProLy 407
QY 1311 CAAGCACACTCTCTACCAGCTGTTATTTGTTTGCAC-----1347
Db 407 sPheArgThrProTrpLeuGlyThrIleLeuLeuGlyValValIleAlaIleAlaLase 427
QY 1348 -----CCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTT 1400
Db 427 rPheLeuProIleSerLeuLeu-----GlyAspLeuValSerLeuGlyThrAl 443
QY 1401 CCTCAGTTTGGCAGGTGCTTTTATTGGCTGGCAGTTGCTGGGCTGATTTATCTTCG 1460
Db 443 aValAlaPheSer-----IleValCysLeuSerValIleTyrLeuAr 457
QY 1461 ATACAAATGCCAGATGATGCATCGTCTCTTCAAGGTGCCA-----CTGTTCCATCCAGC 1514
Db 457 gIleLysHisProAspLeuProArgPropheLysValProGlyIlePheThrAlaAl 477
QY 1515 TTTGTTTTCCTTCACATGCCTCTTC-----ATGGTTGCCCT 1550
Db 477 aAlaGlyIleAlaAlaCysLeuPheLeuProTyrGlnAsnPheGlnProMetIleValHi 497
QY 1551 TTCCCTCTATTGGGACCCATTATTAGTACAGGATTGGCTTCGTCATCACTGACTGGAGT 1610
Db 497 sAlaMetAsnAspAsnProLeuProLeuMetIleLeuGlyGlyTyrAlaAlaValGlyAl 517
QY 1611 CCCTGGGTATTATCTCTTTATATATATGGGACAGAAA 1647
Db 517 aIleIleTyrIleAlaTyrGlyTyrTrpHisserLys 529

```

Search completed: April 16, 2003, 16:26:07
Job time : 65 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 14:34:15 ; Search time 25 Seconds
(without alignments)
7429.240 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_090917_20285/app_query.fasta_1.2375
-DB=SwissProt_40 -QFWT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1.1.24_@runat_04042003_090917_20285 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2519	61.1	501	1 XCT_HUMAN	Q9upy5 homo sapien
2	2307	56.0	502	1 XCT_MOUSE	Q9wtr6 mus musculus
3	1210.5	29.4	507	1 LAT1_HUMAN	Q01650 homo sapien
4	1204	29.2	512	1 LAT1_RAT	Q63016 rattus norv
5	1202	29.2	512	1 LAT1_MOUSE	Q9z127 mus musculus
6	1131.5	27.5	511	1 YLAL1_HUMAN	Q9um01 homo sapien
7	1108	26.9	533	1 LAT2_RAT	Q9wvr6 rattus norv
8	1103	26.8	531	1 LAT2_MOUSE	Q9qxw9 mus musculus
9	1101.5	26.7	535	1 LAT2_HUMAN	Q9uhi5 homo sapien
10	1075.5	26.1	487	1 BAT1_HUMAN	P82251 homo sapien
11	1074.5	26.1	487	1 BAT1_RAT	P82252 rattus norv
12	1067.5	25.9	487	1 BAT1_MOUSE	Q9qxa6 mus musculus
13	368.5	8.9	574	1 MUP1_YEAST	P50276 saccharomyc
14	343.5	8.3	546	1 MUP3_YEAST	P38734 saccharomyc
15	321	7.8	445	1 YHEM_ECOLI	P45539 escherichia
16	300.5	7.3	440	1 YJ99_MYCTU	Q10858 mycobacteri
17	293.5	7.1	658	1 CTR2_HUMAN	P52569 homo sapien
18	286	6.9	629	1 CTR1_HUMAN	P30825 homo sapien

19	279.5	6.8	622	1	CTRL_MOUSE	Q09143 mus musculu
20	275.5	6.7	624	1	CTRL_RAT	P30823 rattus norv
21	273	6.6	657	1	CTR2_MOUSE	P18581 mus musculu
22	272.5	6.6	430	1	YBAT_ECOLI	P77400 escherichia
23	253	6.1	636	1	CTR4_HUMAN	O43246 homo sapien
24	252.5	6.1	459	1	AAPA_BACSU	O06005 bacillus su
25	251.5	6.1	458	1	YDGF_BACSU	P96704 bacillus su
26	250.5	6.1	497	1	ANSP_SALTY	P40812 salmonella
27	239	5.8	489	1	ANSP_MYCTU	O33261 mycobacteri
28	236.5	5.7	499	1	ANSP_ECOLI	P77610 escherichia
29	232	5.6	457	1	PROX_ECOLI	P77327 escherichia
30	232	5.6	467	1	MMUP_ECOLI	Q47689 escherichia
31	232	5.6	550	1	PUFX_EMENI	P18696 emericella
32	228.5	5.5	435	1	Y609_METJA	Q58026 methanococc
33	227.5	5.5	456	1	PROX_SALTY	P37460 salmonella
34	224.5	5.4	481	1	YJ79_MYCTU	Q10875 mycobacteri
35	222.5	5.4	439	1	POTE_ECOLI	P24170 escherichia
36	222.5	5.4	457	1	AROP_ECOLI	P15993 escherichia
37	222.5	5.4	574	1	ME22_SCHPO	O60170 schizosacch
38	222	5.4	573	1	INAI_TRIHA	P34054 trichoderma
39	221	5.4	435	1	POTE_HAEIN	P44768 haemophilus
40	220	5.3	465	1	Y093_RHIME	O87394 rhizobium m
41	218	5.3	611	1	LYP1_YEAST	P32487 saccharomyc
42	216.5	5.3	488	1	LYSP_ECOLI	P25737 escherichia
43	214	5.2	444	1	CADB_ECOLI	P23891 escherichia
44	214	5.2	445	1	YJDE_ECOLI	P39269 escherichia
45	214	5.2	663	1	GNP1_YEAST	P48813 saccharomyc

ALIGNMENTS

RESULT 1					
XCT_HUMAN					
ID	XCT_HUMAN	STANDARD;	PRT;	501 AA.	
AC	Q9UPY5;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Cystine/glutamate transporter (Amino acid transport system xc-) (xCT)				
DE	(Calcium channel blocker resistance protein CCBRI).				
GN	SLC7A11.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21082210; PubMed=11213471;				
RA	Sato H., Tamba M., Kuriyama-Matsumura K., Okuno S., Bannai S.;				
RT	"Molecular cloning and expression of human xCT, the light chain of				
RT	amino acid transport system xc-";				
RL	Antioxid. Redox. Signal. 2:665-671(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Conklin D.S., Beach D.H.;				
RT	"CCBR1, novel CD98 light chain implicated in redox control and calcium				
RT	signaling.";				
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta;				
RA	Chaney C.D., Kekuda R., Wang H., Huang W., Prasad P.D., Smith S.B.,				
RA	Ganapathy V.;				
RT	"Structure, function and regulation of human cystine/glutamate				
RT	transporter in retinal pigment epithelial cells.";				
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Borsani G., Manzoni M., Palacin M., Pineda M., Gasol E.;				
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				

RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY EXCHANGE OF ANIONIC
CC AMINO ACIDS WITH HIGH SPECIFICITY FOR ANIONIC FORM OF CYSTINE AND
CC GLUTAMATE (BY SIMILARITY).
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLC3A2/4F2HC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC -----
CC This SWISS-PROF entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AB026891; BAA82628.1; -
DR EMBL; AF200708; AAG35592.1; -
DR EMBL; AF252872; AAK4911.1; -
DR EMBL; AJ277882; CAC81905.1; -
DR EMBL; BC012087; AAH12087.1; -
DR Genew; HGNC:11059; SLC7A11.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 423 443 POTENTIAL.
FT TRANSMEM 450 470 POTENTIAL.
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 501 AA; 55422 MW; 3EF2648B94A9F59E CRC64;

Alignment Scores:
Pred. No.: 2,22e-190 Length: 501
Score: 2519.00 Matches: 498
Percent Similarity: 96.51% Conservative: 0
Best Local Similarity: 96.51% Mismatches: 0
Query Match: 61.14% Indels: 18
DB: 1 Gaps: 1

US-09-667-170A-440 (1-2239) x XCT_HUMAN (1-501)
QY 184 AAGCCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTAAACGGAGG 243
Db 4 LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlyAsnValAsnGlyArg 23
QY 244 CTGCCTTCCCTGGGCAACAAGGAGGACCTGGGCGAGGACGCTTTTTCAGGAAGAGACGCC 303
Db 24 LeuProSerLeuGlyAsnLysGluProGlyGln----- 35
QY 304 TTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAGTCACT 363
Db 36 -----GluLysValGlnLeuLysArgLysValThr 45
QY 364 TTAAGGGGGAGTCTCCATTATTCATGGCACCATTATGGAGCAGGAATCTTCATCTCT 423
Db 46 LeuLeuArgGlyValSerIleIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 65

QY 424 CCTAAGGGCGTCTCCAGAACACGGGCGAGCGTGGGCAATGTCTCTGACCATCTGGACGGTG 483
Db 66 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal 85
QY 484 TGTGGGTCCTCTCAGTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAG 543
Db 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrThrIleLys 105
QY 544 AAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACACAGCTTTTGT 603
Db 106 LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal 125
QY 604 CGAGTCTGGTGGAACTCTCATATAATACGCCCTGCAGCTACTGTCTGTATATCCCTGGCA 663
Db 126 ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla 145
QY 664 TTTGGACGCTACATCTGGAACCAATTTTATTTCAATGTGAATCCCTGAACCTTGGCATC 723
Db 146 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 165
QY 724 AAGCTCATTACAGCTGTGGGCATAAAGTGTAGTGTGCTTAAATAGCATGAGTGTACG 783
Db 166 LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 185
QY 784 TGGAGCGCCGGATCCAGATTTCTTAAACCTTTTGAAGCTCAGCTCAGCAATTTCTGATAAT 843
Db 186 TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIleIle 205
QY 844 ATAGTCCCTGGAGTTATGACAGCTAATTAAGGTCAAAAGCGAGAACTTTAAAGACGCCCTT 903
Db 206 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 225
QY 904 TCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTATGCA 963
Db 226 SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 245
QY 964 TATGCTGGCTGGTTTACCTCAACTTTTGTACTGAAGAGTAGAAACCCCTGAAAAACC 1023
Db 246 TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 265
QY 1024 ATTCCTCTGTCAATATGATATCCATGGCCATTTGTACCATTTGGCTATGTGCTGACAAAT 1083
Db 266 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 285
QY 1084 GTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTG 1143
Db 286 ValAlaTyrPheThrThrIleAsnAlaGluGluLeuLeuLeuSerAsnAlaValAlaVal 305
QY 1144 ACCTTTTCTGACGGCTACTGGGAATTTCTCATGATTCATGTCGCAAGCACACTCCT 1203
Db 306 ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 325
QY 1204 TCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTTATCTATGTTGCG 1263
Db 326 SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla 345
QY 1264 TCTCGAGAGGGTCACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACTCCT 1323
Db 346 SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro 365
QY 1324 CTACAGCTGTTTATGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACCTC 1383
Db 366 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 385
QY 1384 GACAGCTTTTGAATTTCTCCTCAGTTTTCAGGCTGGCTTTTATTGGGCTGGCAGTTGCT 1443
Db 386 AspSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheIleGlyLeuAlaValAla 405
QY 1444 GGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTCTTCAAGGTGCCACTG 1503
Db 406 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 425
QY 1504 TTCATCCCAGCTTTGTTTTCCTTTCACATGCCCTCTTCATGGTGGCTTCCCTCTATTTCG 1563

QY	784	TGGAGGCCCGGATCCAGATTCTTAAACCTTTTGAAGCTCACAGCAATCTGTATAATT	843
Db	186	TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIle	205
QY	844	ATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACACGAGCAACTTTAAAGACGCTTT	903
Db	206	IleValProGlyValIleGlnLeuIleLysGlyGlnThrHisPheLysAspAlaPhe	225
QY	904	TCAGGACAGATTCAAGTATTACCGGGTTGCCACTGGCTTTTATTATGGAATGATGCA	963
Db	226	SerGlyArgAspThrSerLeuMetGlyLeuProLeuAlaPheTyrTyrGlyMetTyrAla	245
QY	964	TATGCTGGCTGGTTTACCTCACTTGTGTACTGAAGAGTAGAAAAACCTGAAAAAAC	1023
Db	246	TyrAlaGlyTrpPheTyrLeuAsnPheIleThrGluGluValAspAsnProGluLysThr	265
QY	1024	ATTCCTTCTGCAATATGATATCATCGCCCATTTGTCCACCATGGCTATGTGCTGCAAAAT	1083
Db	266	IleProLeuAlaIleCysIleSerMetAlaIleIleThrValGlyTyrValLeuThrAsn	285
QY	1084	GTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCAGTG	1143
Db	286	ValAlaTyrPheThrThrIleSerAlaGluGluLeuLeuGlnSerSerAlaValAlaVal	305
QY	1144	ACCTTTCTGACGGCTACTGGGAAATTTCTCATAGCAGTTCGATCTTTGTTGCCCTC	1203
Db	306	ThrPheSerGluArgLeuLeuGlyLysPheSerLeuAlaValProIlePheValAlaLeu	325
QY	1204	TCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGCTCCAGGTATTCTATGTTGG	1263
Db	326	SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla	345
QY	1264	TCTCGAGAGGCTACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACTCCT	1323
Db	346	SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValHisLysHisThrPro	365
QY	1324	CTACGAGCTGTTATGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACCTC	1383
Db	366	LeuProAlaValIleValLeuHisProLeuThrMetValMetLeuPheSerGlyAspLeu	385
QY	1384	GACAGCTTTTGAATTTCTCAGTTTTCGAGTGGCTTTTATTGGCTGGCAGTTGCT	1443
Db	386	TyrSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheMetGlyLeuAlaValAla	405
QY	1444	GGGCTGATTATCTTCGATACAAATGCCAGATATGATCATGCTCTTCAAGGTGCCACTG	1503
Db	406	GlyLeuIleTyrLeuArgTyrLysArgProAspMetHisArgProPheLysValProLeu	425
QY	1504	TTCATCCAGCTTTGTTTCTTCACATGCCTCTCATGTTTGGCTTTCCTCTCTATTCG	1563
Db	426	PheIleProAlaLeuPheSerPheThrCysLeuPheMetValValLeuSerLeuTyrSer	445
QY	1564	GACCATTTAGTACAGGATGGCTTCGTCATCATCTGACTGGAGTCCCTGGCTATTAT	1623
Db	446	AspProPheSerThrGlyValGlyPheLeuIleThrLeuThrGlyValProAlaTyrTyr	465
QY	1624	CTCTTTATTATATGGACAAAGAACCCAGGTGGTTTACAATAATGTCAGAGAAATAACC	1683
Db	466	LeuPheIleValTrpAspLysLysProLysTrpPheArgArgLeuSerAspArgIleThr	485
QY	1684	AGAACAATACAAATAATACTGGAAGTTGTACCAGAGAGAGATAAG	1728
Db	486	ArgThrLeuGlnIleIleLeuGluValValProGluAspSerLys	500
RESULT 3			
LATI_HUMAN			
ID	LAT1_HUMAN	STANDARD;	PRT; 507 AA.
AC	Q01650; Q9UBN8; Q9UP15; Q9UQC0;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Large neutral amino acids transporter small subunit 1 (L-type amino acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC) (CD98 light		

DE	chain) (Integral membrane protein El6) (hLAT1).
GN	SLC7A5 OR LAT1 OR MPE16 OR CD98LC.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=98421678; PubMed=9751058;
RA	Mastrobardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffing J.,
RA	Shoemaker C.B., Verrey F.;
RT	"Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a
RT	permease family.";
RL	Nature 395:288-291(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Ovary;
RA	Yanagida O., Segawa H., Miyamoto K., Takeda E., Goya T., Endou H.,
RA	Kanai Y.;
RT	"Cloning and characterization of a human system L amino acid
RT	transporter.";
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Placenta;
RX	MEDLINE=99160855; PubMed=10049700;
RA	Prasad P.D., Wang H., Huang W., Kekuda R., Rajan D.P., Leibach F.H.,
RA	Ganapathy V.;
RT	"Human LAT1, a subunit of system L amino acid transporter: molecular
RT	cloning and transport function.";
RL	Biochem. Biophys. Res. Commun. 255:283-288(1999).
RN	[4]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX	MEDLINE=99172172; PubMed=10072483;
RA	Tsurudome M., Ito M., Takebayashi S., Okumura K., Nishio M.,
RA	Kawano M., Kusagawa S., Komada H., Ito Y.;
RT	"Primary structure of the light chain of fusion regulatory protein-
RT	1/CD98/4F2 predicts a protein with multiple transmembrane domains that
RT	is almost identical to the amino acid transporter El6.";
RL	J. Immunol. 162:2462-2466(1999).
RN	[5]
RP	SEQUENCE FROM N.A.
RA	Minato N., Iwai K., Takizawa C., Nakamura E.;
RT	"Human 4F2 light chain: amino acid transporter.";
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN	[6]
RP	SEQUENCE OF 267-507 FROM N.A.
RC	TISSUE=Peripheral blood lymphocytes;
RX	MEDLINE=92283834; PubMed=1597461;
RA	Gaugitsch H.W., Prieschl E.E., Kalthoff F., Huber N.E.,
RA	Baumruker T.;
RT	"A novel transiently expressed, integral membrane protein linked to
RT	cell activation. Molecular cloning via the rapid degradation signal
RT	AUUA.";
RL	J. Biol. Chem. 267:11267-11273(1992).
CC	-!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
CC	NEUTRAL AMINO ACIDS. INVOLVED IN CELLULAR AMINO ACID UPTAKE.
CC	-!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC	TRANSPORT PROTEIN SLC3A2/4F2HC.
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC	-!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN ADULT LUNG AND LIVER,
CC	AND IS ALSO EXPRESSED IN BRAIN, THYMUS, RETINA AND SOME OTHER
CC	TISSUES.
CC	-!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC	PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC	SUBFAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)

or send an email to license@isb-sib.ch).

CC CC EMBL; AF077866; AAC61479.1; -
DR DR EMBL; AB018009; BAA84648.1; -
DR DR EMBL; AF104032; AAD20464.1; -
DR DR EMBL; AB018542; BAA33851.1; -
DR DR EMBL; AB017908; BAA75746.1; -
DR DR EMBL; M80244; AAA35780.1; -
DR DR PIR; A42783; A42783.
DR DR Genew; HGNC:11063; SLC7A5.
DR DR MIM; 600182; -
DR DR InterPro; IPR002293; AA/rel_primeasel.
DR DR InterPro; IPR004760; L_AA_transport.
DR DR InterPro; IPR004841; Permease.
DR DR Pfam; PF00324; aa_permeases; 1.
DR DR TIGRFAMS; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 243 263 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
FT TRANSMEM 319 339 POTENTIAL.
FT TRANSMEM 396 416 POTENTIAL.
FT TRANSMEM 431 451 POTENTIAL.
FT TRANSMEM 458 478 POTENTIAL.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 15 15 A -> V (IN REF. 5).
FT CONFLICT 29 31 AKS -> SKR (IN REF. 5).
FT CONFLICT 35 35 S -> A (IN REF. 5).
FT CONFLICT 62 62 T -> A (IN REF. 5).
FT CONFLICT 88 88 V -> M (IN REF. 5).
FT CONFLICT 154 154 T -> A (IN REF. 5).
FT CONFLICT 230 230 N -> K (IN REF. 1).
SQ SEQUENCE 507 AA; 55010 MW; 767F3C60B62C0F02 CRC64;

Alignment Scores:
Pred. No.: 1.8e-87 Length: 507
Score: 1210.50 Matches: 232
Percent Similarity: 66.19% Conservative: 95
Best Local Similarity: 46.96% Mismatches: 164
Query Match: 29.38% Indels: 3
DB: 1 Gaps: 2

US-09-667-170A-440 (1-2239) x LAT1_HUMAN (1-507)

QY 247 CCTTCCTGGGCAACAAGAGCCACCTGGGCGAGGAGCGCTTTTCAGGAGAGAGCGCCTTT 306
Db 14 ProAlaAlaGluGluLysGluGluAlaArgGluLysMetLeuAlaAlaLysSerAlaAsp 33

QY 307 TCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGCGAAAGTCACTTTA 366
Db 34 GlySerAlaProAlaGlyGluGly--GluGlyValThrLeuGluGlnArgAsnIleThrLeu 52

QY 367 CTGAGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAGATCTTCATCTCTCT 426
Db 53 LeuAsnGlyValAlaIleValGlyThrIleGlySerGlySerGlyIlePheValThrPro 72

QY 427 AAGGGCGTGTCCAGAACACGCGGCGGCGATGTCTCTGACCATCTGGACGGTGTGT 486
Db 73 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValThrAlaAlaCys 92

QY 487 GGGGTCCTGTCACTATTTGGAGCTTTTGTCTTATGCTGAATTTGGAAACAACATAAAGAAA 546
Db 93 GlyValPheSerIleValGlyAlaLeuLysCysTyrAlaGluLeuGlyThrThrIleSerLys 112

QY 547 TCTGGAGGTCAATTACACATATATTTTGAAGTCTTTGGTCCATTACAGCTTTTGTACGA 606
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 492

Db 113 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCTGGGTGGAACTCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCAFTT 666
Db 133 LeuTrpIleGluLeuLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATTTCTGGAACCACTTTTATTCAATGTGAAATCCCTGAACTTGCATCAAG 726
Db 153 AlaThrTyrLeuLeuLysProLeuPheProThrCysProValProGluGluAlaLys 172
QY 727 CTCATTACAGCTGTGGGCATAACTGTAGTGATGGTCTTAATAGCATGAGTGTGCTGG 786
Db 173 LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla 192
QY 787 AGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAATTATA 846
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuIleIle 212
QY 847 GTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCCAGAACTTTAAAGACGCTTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTAT 960
Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
QY 961 GCATATGCTGGTGGTGTACCTCAACTTTTGTACTGAAGAAGTAGAAAACCCCTGAAAAA 1020
Db 253 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACCATTCCCTTGCAT 1080
Db 273 AsnLeuProLeuAlaIleIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
QY 1081 AATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAATGCAGTGGCA 1140
Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATATTAGCAGTTCCGATCTTTGTTGCC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
QY 1201 CTCTCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGTGCTGCTCCAGGTTATTTCTATGTT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPhePheVal 352
QY 1261 GCGTCTCGAGAGGTCACCTTCCAGAAATCTCTCCATGATTGATGTCGCAAGCACACT 1320
Db 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
QY 1321 CCTCTACAGCTGTTATTGTTTGCACCTTTGACAAATGATAATGCTCTTCTCTGGAGAC 1380
Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
QY 1381 CTCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTT 1440
Db 393 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 412
QY 1441 GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTTCAAGGTGCCA 1500
Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
QY 1501 CTGTTTCATCCAGCTTTGTTTCTTCATGCTCTTCATGGTGGCTTTCCCTCTAT 1560
Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452
QY 1561 TCGGACCATTTAGTACAGGAGTTGGCTTCGTTCATCTACTCTGAGTGGAGTCCCTGCGTAT 1620
Db 453 LysThrProValGluCysGlyIleGlyPheThrIleLeuSerGlyLeuProValTyr 472
QY 1621 TATCTCTTATATATGGGCAAGAAACCCAGGTGGTTTAGAATAATGTCAGAGAAAAATA 1680
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 492

111 IleSerLysSerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAla 130
 598 TTTGTACGAGTCTGGTGGAACTCCTCATATAACGCCCTGCAGCTACTGCTGTGATATCC 657
 131 PheLeuLysLeuTrpIleGluLeuLeuIleArgProSerSerGlnTyrIleValAla 150
 658 CTGGCATTTGGACGCTACATCTTGGAAACCATTTTTTATCAATGTGAAATCCCTGAACTT 717
 151 LeuValPheAlaThrTyrLeuLeuLysProValPheProThrCysProValProGluGlu 170
 718 GCGATCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGGTCTCTAAATAGCATGAGT 777
 171 AlaAlaLysLeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSer 190
 778 GTCAGCTGGAGCGCCCGGATCCAGATTTCTTAACCTTTTGGCAAGCTCACAGCAATTCGT 837
 191 ValLysAlaAlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAla 210
 838 ATAATTATAGTCCCTGGAGTTATGCAGCTAATTAATA-----GGTCAACAGCAG 885
 211 LeuIleIleLeuLeuGlyPheIleGlnMetGlyLysAspMetGlyGlnGlyAspAlaSer 230
 886 AACTTT-----AAAGACGCCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTG 939
 231 AsnLeuGlnGlnLysLeuSerPheGluGlyThrAsnLeuAspValGlyAsnIleValLeu 250
 940 GCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAA 999
 251 AlaLeuTyrSerGlyLeuPheAlaTyrGlyTyrPAsnTyrLeuAsnPheValThrGlu 270
 1000 GAAGTAGAAAACCTGAAAAAACCATTCCTTGCATATGTATATCCATGGCCATTGTC 1059
 271 GluMetIleAsnProTyrArgAsnLeuProLeuAlaIleIleSerLeuProIleVal 290
 1060 ACCATTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTG 1119
 291 ThrLeuValTyrValLeuThrAsnLeuAlaTyrPheThrThrLeuSerThrAsnGlnMet 310
 1120 CTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTA 1179
 311 LeuThrSerGluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrp 330
 1180 GCAGTTCCGATCTTTGTGCTGCTCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCT 1239
 331 IleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThr 350
 1240 GTCTCCAGGTTATTCTATGTGCGTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATG 1299
 351 SerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProSerValLeuSerMet 370
 1300 ATTCAATGTCGCAAGCACACTCTCTACCAAGCTGTTATTTGTCACCCCTTTTGACAATG 1359
 371 IleHisProGlnLeuLeuThrProValProSerLeuValPheThrCysIleMetThrLeu 390
 1360 ATAATGCTCTTCTCTGGAGACTCGACAGTCTTTTGAATTTCTCTCAGTTTGGCCAGTGG 1419
 391 MetTyrAlaPheSerArgAspIlePheSerIleAsnPhePheSerPhePheAsnTrp 410
 1420 CTTTTTATGGGCTGGCAGTTGCTGGGCTGATTTATCTTCATACAAATGCCAGATATG 1479
 411 LeuCysValAlaLeuAlaIleIleGlyMetMetTrpLeuArgPheLysLysProGluLeu 430
 1480 CATCGTCTTTCAAGGTGGCAGTGTCTATCCATCCAGCTTTTGTTCCTTCACATGCCTCTTC 1539
 431 GluArgProIleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPhe 450
 1540 ATGGTTGCCCTTTCCTCTATTCGGACCCATTTAGTACAGGGAFTGGCTTCGTCAATCACT 1599
 451 LeuIleAlaValSerPheTrpLysThrProMetGluCysGlyIleGlyPheAlaIle 470
 1600 CTGACTGGAGTCCCTGCGTATTATCTCTTATTATATGGGACAGAAACCCAGGTGGTTT 1659


```
QY 850 CCTGGAGTATGACGCTAATTAAGGTCAACGCGAGAACTTTAAAGACGCCTTTTCAGGA 909
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
201 AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly 220
QY 910 AGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGATGATGCT 969
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
221 SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 240
QY 970 GGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAACCATTC 1029
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
241 GlyTrpAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro 260
QY 1030 CTTGCAATATGTATATCCATGGCCATGTCCACCATGGCTATGTCTGCAAAATGTGG 1089
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
261 LeuSerIleGlyIleSerMetProIleValThrIleIleTyrIleLeuThrAsnValAla 280
QY 1090 TACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCAGCTTT 1149
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
281 TyrTyrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe 300
QY 1150 TCTGAGCGGCTACTGGAAATTTCTCATATTAGCAGTTCGGAFTTGTGGCTCTCCTGC 1209
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
301 AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys 320
QY 1210 TTTGGCTCCATGACGGTGGTGTGTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCG 1269
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
321 PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuPhePheValGlySerArg 340
QY 1270 GAGGTCACCTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCACACTCCTCTACCA 1329
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
341 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 360
QY 1330 GCTGTATTGTTTGCACCTTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGT 1389
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
361 SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln 380
QY 1390 CTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGTGGCAGTGTCTGGGCTG 1449
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
381 LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln 400
QY 1450 ATTTATCTCGATACAAATGCCAGATATGTCATCGTCCTTTCAAGGTGCCACTGTTTCATC 1509
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
401 LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe 420
QY 1510 CCAGCTTTGTTTCTTCCATGCTCTTCCATGCTGCTTCCCTCTATTCGGACCCA 1569
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
421 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr 440
QY 1570 TTTAGTACAGGATGGCTTCGTCATCTGCTGAGTGGCTCCCTGCGTATATCTCTTT 1629
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle 460
QY 1630 ATT-----ATATGGGACAAAGAACCCAGGTGGTTTAGATAATATGTCAGAGAAATAACC 1683
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
461 IleArgValProGluHisLysArgProLeuTyrLeuArgArgIleValGlySerAlaThr 480
QY 1684 AGAACATTACAATAATA---CTGGAAGTTGTACCAGAGAGAGAT 1725
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
481 ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp 495
RESULT 7
LAT2_RAT
ID LAT2_RAT STANDARD; PRT; 533 AA.
AC Q9WVR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large neutral amino acids transporter small subunit 2 (L-type amino
DE acid transporter 2).
GN SLC7A8 OR LAT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

```
OX NCBI_TaxID=10116;
.RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99321902; PubMed=10391916;
RX Segawa H., Fukasawa Y., Miyamoto K., Takeda E., Endou H., Kanai Y.;
RA "Identification and functional characterization of a Na+-independent
RT neutral amino acid transporter with broad substrate selectivity.";
RL J. Biol. Chem. 274:19745-19751(1999).
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
CC NEUTRAL AMINO ACIDS. HAS HIGHER AFFINITY FOR L-PHENYLALANINE THAN
CC LAT1. L-ALANINE IS TRANSPORTED AT PHYSIOLOGICAL CONCENTRATIONS.
CC PLAYS A ROLE IN BASOLATERAL (RE)ABSORPTION OF NEUTRAL AMINO ACIDS.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB024400; BAA82517.1; -.
CC InterPro; IPR002293; AA/rel_primease1.
CC InterPro; IPR004760; L_AA_transport.
CC InterPro; IPR004841; Permease.
CC Pfam; PF00324; aa_permeases; 1.
CC TIGRFAMs; TIGR00911; 2A0308; 1.
CC Transport; Amino-acid transport; Transmembrane.
CC FT TRANSMEM 41 61 POTENTIAL.
CC FT TRANSMEM 73 93 POTENTIAL.
CC FT TRANSMEM 114 134 POTENTIAL.
CC FT TRANSMEM 156 176 POTENTIAL.
CC FT TRANSMEM 190 210 POTENTIAL.
CC FT TRANSMEM 232 252 POTENTIAL.
CC FT TRANSMEM 269 289 POTENTIAL.
CC FT TRANSMEM 311 331 POTENTIAL.
CC FT TRANSMEM 363 383 POTENTIAL.
CC FT TRANSMEM 389 409 POTENTIAL.
CC FT TRANSMEM 423 443 POTENTIAL.
CC FT TRANSMEM 448 468 POTENTIAL.
CC SEQUENCE 533 AA; 58190 MW; 99479DB60DA69DF0 CRC64;
SQ
```

Alignment Scores: 2.09e-79 Length: 533
Pred. NO.: 1108.00 Matches: 224
Score: 63.04% Conservative: 100
Percent Similarity: 43.58% Mismatches: 168
Best Local Similarity: 26.89% Indels: 22
Query Match: 1 Gaps: 5
DB:

US-09-667-170A-440 (1-2239) x LAT2_RAT (1-533)
QY 202 ATCTCCAAAGGAGGTTACCTGCAGGGGAAATGTTAACGGGAGGCTGCCTTCCCTGGGCAAC 261
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 20
1 MetGluLysGlyThrArgGlnArgAsnAsnThrAlaLysAsnHisProAspArgGlySer 20
QY 262 AAGGAGCCACCTGGGACGAGCGCTTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGACGCC 321
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 34
21 AspThrSerProGluAlaGluAlaSerSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 34
QY 322 TTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAGAAAGTCACTTTTACTGAGGGAGTCTCC 381
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 48
35 -----ValAlaLeuLysLysGluIleGlyLeuValSerAlaCysGly 48
QY 382 ATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCATCTCTCTCTAAGGGCGTCTCCAG 441
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 441

Db 49 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 68
QY 442 AACACGGGCGAGCGTGGCATGTCTCTGACCATCTGGACGGTGTGGGGTCCCTGCTCACTA 501
Db 69 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyValIleThrAla 88
QY 502 TTTGGAGCTTTGCTTATGCTGAATGGGAACAACATATAAGAAATCTGGAGGTCAATAC 561
Db 89 ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 108
QY 562 ACATATAATTTGGAAGTCTTTGGTCCATTACCAAGCTTTGTACGAGTCTGGTGGAACTC 621
Db 109 SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 128
QY 622 CTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGCTACATCTG 681
Db 129 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 148
QY 682 GAACCATTTTATTCATCAATCTGAATCCCTGAACTCCGATCGATCAAGCTCATTACAGCTGTG 741
Db 149 GlnProLeuPheProThrCysPheProProGluSerGlyLeuArgLeuAlaIle 168
QY 742 GGCATAACTGTAGTGATGCTCTAAATAGCATGAGTGTGACGTGGACGCCGGATCCAG 801
Db 169 CysLeuLeuLeuLeuThrTrpValAsnCysSerValArgTrpAlaThrArgValGln 188
QY 802 ATTTCTTAACTTTGCAAGCTCACAGCAATCTGATAATATAGTCCCTGGAGTTATG 861
Db 189 AspIlePheThrAlaGlyLysLeuAlaLeuAlaLeuIleIleMetGlyValVal 208
QY 862 CAGCTAATTAAGGTCAACACGCAGAACTTT-----AAAGACGCCCTTTTCA 906
Db 209 GlnIleCysLysGlyGlu-----PhePheTrpLeuGluProLysAsnAlaPheGlu 225
QY 907 GGA---AGAGATCAAGTATTACCGGGTGGCCACTGGCTTTTATTATGGAATGATGCA 963
Db 226 AsnPheGlnGluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAla 245
QY 964 TATGCTGGCTGGTTTACCTCAACTTGTACTGAAGAGTAGAAACCCCTGAAACAAACC 1023
Db 246 TyrGlyGlyTrpAsnPheLeuAsnTyrValThrGluLeuValAspProTyrLysAsn 265
QY 1024 ATCCCTTCTGCAATATGATATCATCATGCGCATTTGTCACCATGGCTATGCTGACAAAT 1083
Db 266 LeuProArgAlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsn 285
QY 1084 GTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAATGCAAGTGGCAGTG 1143
Db 286 IleAlaTyrValThrAlaMetSerProGlnGluLeuAlaSerAsnAlaValAlaVal 305
QY 1144 ACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATATAGCAGTTCCGATCTTTGTCCTC 1203
Db 306 ThrPheGlyGluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeu 325
QY 1204 TCCTGCTTTGGCTCCATCAACGGTGGTGTGTTGCTGTCTCCAGTTATTCTATGTTGCG 1263
Db 326 SerThrPheGlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGly 345
QY 1264 TCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGCAAGCACACTCT 1323
Db 346 AlaArgGluGlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrPro 365
QY 1324 CTACCAAGCTGTTATGTTTGCACCCCTTTGCAATGATAATGCTCTCTCTGGAGACCTC 1383
Db 366 IleProAlaLeuLeuPheThrCysLeuSerThrLeuLeuMetLeuValThrSerAspMet 385
QY 1384 GACAGTCTTTGAAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGCTGGCAGTTGCT 1443
Db 386 TyrThrLeuIleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAla 405
QY 1444 GGGCTGATTATCTTCGATACAAATGCCAGATATGTCATGCTCTTCAAGGTGCCACTG 1503
Db 406 GlyGlnIleValLeuArgTrpLysLysProAspIleProArgProIleLysIleSerLeu 425

QY 1504 TTCATCCCAGCTTTGTTTTCCTTCCATGCCTTTCATGTTGCTTCCCTTCCCTCTATTTCG 1563
Db 426 LeuPheProIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 445
QY 1564 GACCCATTTAGTACAGGATTTGGCTTTCGTCATCTGACTGGAGTCCCTGCGTATTAT 1623
Db 446 GluProValValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPhe 465
QY 1624 CTCTTTTATTATATGGACAAAGAAACCCAGGTGGTTTGAATAATGTCAGAGAAATAACC 1683
Db 466 LeuGlyValTyrTrpGlnHisLysProLysCysPheAsnAspPheIleGluSerLeuThr 485
QY 1684 AGAACATTACAAATAATACTGGAAGTTGTA---CCAGAAGAA 1722
Db 486 LeuValSerGlnLysMetCysValValValValTyrProGlnGlu 499

RESULT 8

LAT2_MOUSE
ID LAT2_MOUSE STANDARD; PRT; 531 AA.
AC Q9QXW9;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large neutral amino acids transporter small subunit 2 (L-type amino acid transporter 2).
GN SLC7A8 OR LAT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=20044753; PubMed=10574970;
RA Rossier G., Meier C., Bauch C., Summa V., Sordat B., Verrey F., Kuhn L.C.;
RT "LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter of kidney and intestine.";
RL J. Biol. Chem. 274:34948-34954(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079165; PubMed=10610726;
RA Bassi M.T., Sperandeo M.P., Incerti B., Bulfone A., Pepe A., Surace E.M., Gattuso C., de Grandi A., Buoinconti A., Riboni M., Manzoni M., Andria G., Ballabio A., Borsani G., Sebastio G.;
RT "SLC7A8, a gene mapping within the lysinuric protein intolerance critical region, encodes a new member of the glycoprotein-associated amino acid transporter family.";
RL Genomics 62:297-303(1999).
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE NEUTRAL AMINO ACIDS. HAS HIGHER AFFINITY FOR L-PHENYLALANINE THAN LAT1. L-ALANINE IS TRANSPORTED AT PHYSIOLOGICAL CONCENTRATIONS. PLAYS A ROLE IN BASOLATERAL (RE)ABSORPTION OF NEUTRAL AMINO ACIDS.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: MAINLY EXPRESSED IN KIDNEY AND INTESTINE.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT) SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF171668; AAF20380.1; -.
CC EMBL; Y19022; CAB69072.1; -.
CC MGI; 1355323; Slc7a8.
CC -----
CC DR
CC DR
CC DR

DR InterPro: IPR002293; AA/rel_primease1.
DR InterPro: IPR004760; L_AA_transport.
DR InterPro: IPR004841; Permease.
DR Pfam: PF00324; aa_permeases; 1.
DR TIGRFAMs: TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
FT TRANSMEM 309 329 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 387 407 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
FT TRANSMEM 446 466 POTENTIAL.
SQ SEQUENCE 531 AA; 57873 MW; AE9C3B42F3B24F8C CRC64;

Alignment Scores: 5.17e-79 Length: 531
Pred. NO.: 1103.00 Matches: 220
Score: 64.57% Conservative: 99
Percent Similarity: 44.53% Mismatches: 159
Best Local Similarity: 26.77% Indels: 16
Query Match: 1 Gaps: 5
DB:

US-09-667-170A-440 (1-2239) x LAT2_MOUSE (1-531)

QY	262	AAGGACCCACCTGGGAGGAGCGCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGACGCGC	321
Db	13	LysAsnHisProGlySerAspThrSerProGluAlaGluAlaSerSerGlyGlyGly 32	
QY	322	TTTTTCAGGAAGAGAGAACTGTCAGCTGAAGAGGAAAGTCACTTTTACTGAGGGGAGTCTCC	381
Db	33	-----ValAlaLeuLysLysLysGluLeuGlyLeuValSerAlaCysGly 46	
QY	382	ATTATCATTTGGCACCACCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCCAG	441
Db	47	IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysProLysValLeuGlu 66	
QY	442	AACACGGGAGCGTGGGCATGTCTGTGACCATCTGGACGGTGTGTGGGTCCTGTCACTA	501
Db	67	AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyIleIleThrAla 86	
QY	502	TTTGGAGCTTTGTCTTATGTCTGAATTTGGGAACAACATAAAGAAATCTGGAGTCAATAC	561
Db	87	ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 106	
QY	562	ACATATATTTTGAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGTGGGAACTC	621
Db	107	SerTyrValLysAspIlePheGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 126	
QY	622	CTCATAATACCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATCTG	681
Db	127	LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 146	
QY	682	GAACCATTTTATTAATCAATGTGAATCCCTGAACTTGGCATCAAGCTCATTTACAGCTGTG	741
Db	147	GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuLeuAlaIle 166	
QY	742	GGCAAACTGTAGTATGGTCTCTAATAGCATGAGTGTGAGCTGGAGCGCCCGGATCCAG	801
Db	167	CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 186	
QY	802	ATTTTCTTAACCTTTTGCAGAGCTCACAGCAATCTGTGATAATTATAGTCCCTGGAGTTATG	861
Db	187	AspIlePheThrAlaGlyLysLeuLeuAlaLeuAlaLeuIleIleMetGlyIleVal 206	
QY	862	CAGCTAATTAAGGTCAACGACGAGAACTTT-----AAAGACGCTTTTCA	906
Db	207	GlnIleCysLysGlyGlu-----phePheTrpLeuGluProLysAsnAlaPheGlu 223	

QY	907	GGA---AGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATGCA	963
Db	224	AsnPheGlnGluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAla	243
QY	964	TATGCTGGCTGGTTCCTCACTCAACTTGTACTGAAGAAGTAGAAAACCCCTGAAAAAAC	1023
Db	244	TyrGlyGlyTrpAsnPheLeuAsnTyrValThrGluLeuValAspProTyrLysAsn	263
QY	1024	ATTCCCTTGTCAATATGATATATCCATGGCCATTGTACCATTTGGCTATGTGCTGACAAAT	1083
Db	264	LeuProArgAlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsn	283
QY	1084	GTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCAGTG	1143
Db	284	IleAlaTyrValThrAlaMetSerProGlnGluLeuAlaSerAsnAlaValAlaVal	303
QY	1144	ACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATATTAGCAGTTCCGATCTTTGTTGCCCTC	1203
Db	304	ThrPheGlyGluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeu	323
QY	1204	TCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGTTATTTCTATGTTGCG	1263
Db	324	SerThrPheGlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGly	343
QY	1264	TCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTCCT	1323
Db	344	AlaArgGluGlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrPro	363
QY	1324	CTACGAGCTGTTATTTGTTGTCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTC	1383
Db	364	IleProAlaLeuLeuPheThrCysLeuSerThrLeuLeuMetLeuValThrSerAspMet	383
QY	1384	GACAGCTTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCT	1443
Db	384	TyrThrLeuIleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAla	403
QY	1444	GGGCTGATTTATTCGATACAAATGCCAGATATGATGCTGCTCTTCAAGTGCCACTG	1503
Db	404	GlyGlnIleValLeuArgTrpLysLysProAspIleProArgProIleLysValSerLeu	423
QY	1504	TTCATCCCAGCTTTGTTTCCCTTCACATGCTCTTCTCATGGTGGCCCTTTCCCTCTATTG	1563
Db	424	LeuPheProIleIleTyrLeuLeuPheThrAlaPheLeuLeuIlePheSerLeuTrpSer	443
QY	1564	GACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTAT	1623
Db	444	GluProValValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPhe	463
QY	1624	CTCTTTTATATATGGACAAGAAACCCAGCTGGTGTAGAAATAATGTAGAGAAATAACC	1683
Db	464	LeuGlyValTyrTrpGlnHisLysProLysCysPheAsnAspPheIleLysSerLeuThr	483
QY	1684	AGAACATTACAAATAATACTGGAAGTTGTA---CCAGAAGAA 1722	
Db	484	LeuValSerGlnLysMetCysValValValTyrProGlnGlu 497	

RESULT 9
LAT2_HUMAN
ID LAT2_HUMAN STANDARD; PRT; 535 AA.
AC Q9UHI5; Q9Y445; Q9UKQ6; Q9UKQ7; Q9UKQ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Large neutral amino acids transporter small subunit 2 (L-type amino
DE acid transporter 2) (hLAT2).
GN SLC7A8 OR LAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

QY 1033 GCAATATGATATCCATGGCCATTTGTCACCAATGGCTATGTGCTGACAAATGTGGCCTAC 1092
||||| ||||||| : : : ||||| ||||| ||||||| |||||
Db 268 AlallePheIleSerIleProLeuValThrPheValThrPheAlaAsnValAlaTyr 287
||| : : : ||||||| : : : ||||| ||||| ||||||| |||||
QY 1093 TTTACCAACATTAATGCTGAGGAGCTGCTGCTTTTCAATGCAGTGGCAGTGACCTTTTCT 1152
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 288 ValThrAlaMetSerProGlnGluLeuAlaSerAsnAlaValAlaValThrPheGly 307
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1153 GAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCTGCTTTT 1212
||| : : : ||||||| : : : ||||||| ||||||| ||||||| |||||
Db 308 GluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeuSerThrPhe 327
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1213 GGCTCCATGAACGGTGGTGTGTTGCTGCTCTCCAGTTATCTATGTTGCTGCTCGAGAG 1272
||| : : : ||||||| : : : ||||||| ||||||| ||||||| |||||
Db 328 GlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGlyAlaArgGlu 347
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1273 GGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGAAGCACACTCCTCTACAGCT 1332
||| : : : ||||||| : : : ||||||| ||||||| ||||||| |||||
Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrProIleProAla 367
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1333 GTTATTGTTTGCACCCCTTTGACAAATGATAATGCTCTCTCTGAGACCTCGACAGCTTT 1392
||| : : : ||||||| : : : ||||||| ||||||| ||||||| |||||
Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu 387
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1393 TTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATT 1452
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 388 IleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle 407
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1453 TATCTTCGATACAAATGCCAGATATGATGATGCTCTCTTCAAGTGCCACTGTTTCATCCA 1512
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 408 ValLeuArgTrpLysLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1513 GCTTTGTTTCTCATGCTCTCTCATGGTTGCCCTTTTCCCTCTATTCGGACCCATT 1572
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 428 IleIleTyrLeuLeuPheTrpAlaPheLeuLeuValPheSerLeuTrpSerGluProVal 447
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1573 AGTACAGGATTTGGCTCGTCATCACTCTGCTGAGTCCCTGCGTATTCCTCTTTATT 1632
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 448 ValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValThrPheLeuGlyVal 467
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1633 ATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTGAGAAATAACAGAACATTA 1692
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuLeuThrLeuValSer 487
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| |||||
QY 1693 CAAATAATACTGGAAGTTGTA---CCAGAAGAGATAAG 1728
||| : : : ||||||| ||||||| ||||||| : : : : :
Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500
||| : : : ||||||| ||||||| ||||||| : : : : :
RESULT 10
BAT1_HUMAN
ID BAT1_HUMAN STANDARD; PRT; 487 AA.
AC P82251;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B(0,+)-type amino acid transporter 1 (B(0,+AT) (Glycoprotein-
associated amino acid transporter b0,+AT1).
GN SLC7A9 OR BAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., VARIANTS CSNU3 ARG-105; MET-170; THR-182; ARG-195
AND ARG-259, MUTAGENESIS OF VAL-170, FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Kidney;
RX MEDLINE=99400552; PubMed=10471498;
RA Feliubadalo L., Font M., Purroy J., Rousaud F., Estivill X., Nunes V.,
Golomb E., Centola M., Aksentijevich I., Kreiss Y., Goldman B.,
Pras M., Kastner D.L., Pras E., Gasparini P., Bisceglia L., Beccia E.,
Gallucci M., De Sanctis L., Ponzone A., Rizzoni G.F., Zelante L.,
Bassi M.T., George A.L. Jr., Manzoni M., De Grandi A., Riboni M.,

RA Endsley J.K., Ballabio A., Borsani G., Reig N., Fernandez E.,
RA Estevez R., Pineda M., Torrents D., Camps M., Lloberas J., Zorzano A.,
RA Palacin M.;
RT "Non-type I cystinuria caused by mutations in SLC7A9, encoding a
subunit (b0,+AT) of rBAT.";
RL Nat. Genet. 23:52-57(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=20056095; PubMed=10588648;
RA Pfeiffer R., Loffing J., Rossler G., Bauch C., Meier C., Eggermann T.,
RA Loffing-Cueni D., Kuehn L.C., Verrey F.;
RT "Luminal heterodimeric amino acid transporter defective in
cystinuria.";
RL Mol. Biol. Cell 10:4135-4147(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=21219147; PubMed=11318953;
RA Mizoguchi K., Cha S.H., Chairoungdua A., Kim J.Y., Shigeta Y.,
RA Matsuo H., Fukushima J., Awa Y., Akakura K., Goya T., Ito H.,
RA Endou H., Kanai Y.;
RT "Human cystinuria-related transporter: localization and functional
characterization.";
RL Kidney Int. 59:1821-1833(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE HIGH-AFFINITY, SODIUM-INDEPENDENT
TRANSPORT OF CYSTINE AND NEUTRAL AND DIBASIC AMINO ACIDS (SYSTEM
B(0,+)-LIKE ACTIVITY). THOUGHT TO BE RESPONSIBLE FOR THE HIGH-
AFFINITY REABSORPTION OF CYSTINE IN THE KIDNEY TUBULE.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
TRANSPORT PROTEIN SLC3A1.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: KIDNEY, SMALL INTESTINE, LIVER AND PLACENTA.
CC -!- DISEASE: DEFECTS IN SLC7A9 ARE A CAUSE OF CYSTINURIA WHICH ARISES
FROM IMPAIRED TRANSPORT OF CYSTINE AND DIBASIC AMINO ACIDS THROUGH
THE EPITHELIAL CELLS OF THE RENAL TUBULE AND GASTROINTESTINAL
TRACT. THREE TYPES OF CYSTINURIA HAVE BEEN DESCRIBED. DEFECTS IN
SLC7A9 ARE ASSOCIATED WITH TYPE III CYSTINURIA (CSNU3).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF141289; AAD55898.1; -;
EMBL; AJ249199; CAB54003.1; -;
EMBL; AB033548; BAB16840.1; -;
EMBL; BC017962; AAH17962.1; -;
Genew; HGNC:11067; SLC7A9.
MIM; 604144; -;
MIM; 600918; -;
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa-permeases; 1.
DR Transport; Amino-acid transport; POTENTIAL.
KW TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 252 272 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 349 369 POTENTIAL.

FT	TRANSMEM	375	395	POTENTIAL.
FT	TRANSMEM	410	430	POTENTIAL.
FT	TRANSMEM	435	455	POTENTIAL.
FT	VARIANT	105	105	G -> R (IN CSNU3).
FT	VARIANT	170	170	/FTID=VAR_010256.
FT	VARIANT	182	182	V -> M (IN CSNU3).
FT	VARIANT	195	195	/FTID=VAR_010257.
FT	VARIANT	259	259	A -> T (IN CSNU3).
FT	VARIANT	170	170	/FTID=VAR_010258.
FT	VARIANT	259	259	G -> R (IN CSNU3).
FT	VARIANT	170	170	/FTID=VAR_010259.
FT	MUTAGEN	52	52	/FTID=VAR_010260.
FT	CONFLICT	52	52	V->M: LOSS OF AMINO ACID TRANSPORT
FT	SEQUENCE	487 AA;	53481 MW;	ACTIVITY.
SQ				P -> S (IN REF. 2).
				EF2C30DDE15594F1 CRC64;
Alignment Scores:				
Pred. No.:	7.43e-77	Length:	487	
Score:	1075.50	Matches:	207	
Percent Similarity:	66.16%	Conservative:	98	
Best Local Similarity:	44.90%	Mismatches:	153	
Query Match:	26.10%	Indels:	3	
DB:	1	Gaps:	2	
US-09-667-170A-440 (1-2239) x BAT1_HUMAN (1-487)				
QY	346	CTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGA	405	
Db	26	LeuGlnLysGluLeuGlyLeuIleSerGlyIleSerIleValGlyThrIleGly	45	
QY	406	GCAGGAATCTTCATCTCTCCTAAGGCGGTGCTCCAGAACACAGCGGTGGCATGTCT	465	
Db	46	SerGlyIlePheValSerProLysSerValLeuSerAsnThrGluAlaValGlyProCys	65	
QY	466	CTGACCATCTGGACGGTGTGGGGTCTGTCACATATTGGAGCTTTGTCTTATGCTGAA	525	
Db	66	LeuIleIleTrpAlaAlaCysGlyValLeuAlaThrLeuGlyAlaLeuCysPheAlaGlu	85	
QY	526	TTGGGAACAACATAAAGAAATCTGGAGGTCAATTACACATATATTTTGGAGTCTTTGGT	585	
Db	86	LeuGlyThrMetIleThrLysSerGlyGlyGluProTyrLeuMetGluAlaTyrGly	105	
QY	586	CCATTACCAGCTTTTGTACAGTCTGGGTGGAACCTCTCATATAACGSCCTGCAGCTACT	645	
Db	106	ProIleProAlaTyrLeuPheSerTrpAlaSerLeuIleValIleLysProThrSerPhe	125	
QY	646	GCTGTGATATCCCTGGCAATTTGGACGCTACATCTTGGAAACCATTTTATTCAATGTGAA	705	
Db	126	AlaIleIleCysLeuSerPheSerGluTyrValCysAlaProPheTyrValGlyCysLys	145	
QY	706	ATCCCTGAATCGGATCAAGCTCATTACAGCTGTGGGCATCACTGTAGTGTGCTCA	765	
Db	146	ProProGlnIleValLysCysLeuAlaAlaAlaIleLeuPheIleSerThrVal	165	
QY	766	AATAGCATGAGTGTCAAGCTGGAGCGCCGATCCAGATTTTCTTAACCTTTTGCAGCTC	825	
Db	166	AsnSerLeuSerValArgLeuGlySerTyrValGlnAsnIlePheThrAlaAlaLysLeu	185	
QY	826	ACAGCAATCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAACGCAG	885	
Db	186	ValIleValAlaIleIleIleSerGlyLeuValLeuAlaGlnGlyAsnThrLys	205	
QY	886	AACCTTAAAGACGCTTTTCAGGAGAGATTCAGATTAACGGTGGTGGCTGCTTTT	945	
Db	206	AsnPheAspAsnSerPheGluGlyAlaGlnLeuSerValGlyAlaIleSerLeuAlaPhe	225	
QY	946	TATTATGGAATGATGATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTA	1005	
Db	226	TyrAsnGlyLeuTrpAlaTyrAspGlyTrpAsnGlnLeuAsnTyrIleThrGluGluLeu	245	
QY	1006	GAACACCTGAAACCAATTCCTTCCCTTGCAATATGATATCATTCATGGCCATTTGTCACCAT	1065	

Db	246	ArgAsnProTyrArgAsnLeuProLeuAlaIleIleGlyIleProLeuValThrAla	265	
QY	1066	GGCTATGTGCTGACAAATGTGGCTTACGACCATTAATGCTGAGGAGCTGCTGCTT	1125	
Db	266	CysTyrIleLeuMetAsnValSerTyrPheThrValMetThrAlaThrGluLeuGln	285	
QY	1126	TCAAATGCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTT	1185	
Db	286	SerGlnAlaValAlaValThrPheGlyAspArgValLeuTyrProAlaSerTrpIleVal	305	
QY	1186	CCGATCTTTGTTGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCC	1245	
Db	306	ProLeuPheValAlaPheSerThrIleGlyAlaAlaAsnGlyThrCysPheThrAlaGly	325	
QY	1246	AGGTATTCTTATGCTGCTCGAGAGGTGACCTTCCAGAAATCCTCTCCATGATTCAT	1305	
Db	326	ArgLeuIleTyrValAlaGlyArgGluGlyHisMetLeuLysValLeuSerTyrIleSer	345	
QY	1306	GTCCGCAACACACTCCTCTACCAGCTGTTATTGTTTGCACCCTTTTGACAAATGATAATG	1365	
Db	346	ValArgArgLeuThrProAlaProAlaIleIlePheTyrGlyIleAlaThrIleTyr	365	
QY	1366	CTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTT	1425	
Db	366	IleIleProGlyAspIleAsnSerLeuValAsnTyrPheSerPheAlaAlaTrpLeuPhe	385	
QY	1426	ATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGT	1485	
Db	386	TyrGlyLeuThrIleLeuGlyLeuIleValMetArgPheThrArgLysGluLeuGluArg	405	
QY	1486	CCTTCAAGGTGCCACTGTTTCCAGCTTTTGTGTTTCCATCATGCTCTTTCATGGTT	1545	
Db	406	ProIleLysValProValIleProValLeuMetThrLeuIleSerValPheLeuVal	425	
QY	1546	GCCCTTCTCCTCTATTTCGGACCCATTTAGTACAGGATTTGGCTTCGTCATC---	1602	
Db	426	LeuAlaProIleIleSerLysProThrTrpGluTyrLeuTyrCysValLeuPheIleLeu	445	
QY	1603	ACTGGAGTCCCTGCTATTATCTCTTTATATATATGGACAAAGAACCCAGGTGTTTGA	1662	
Db	446	SerGlyLeuLeuPheTyrPheLeuPheValHisTyr-----LysPheGlyTrpAlaGln	463	
QY	1663	ATAATGTCAGAGAAATAACAGAACATTAATAATACTGGAAGTTGTACCAGAGAA	1722	
Db	464	LysIleSerLysProIleThrMetHisLeuGlnMetLeuMetGluValValProGlu	483	
QY	1723	GAT 1725		
Db	484	Glu 484		

RESULT 11
BAT1_RAT
ID BAT1_RAT
AC P82252;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B(0,+)-type amino acid transporter 1 (B(0,+)-AT) (Glycoprotein-
associated amino acid transporter b0,+AT1).
GN SLC7A9 OR BAT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, SUBUNITS, AND
RP SUBCELLULAR LOCATION.
RC TISSUE=Kidney;
RX MEDLINE=99436074; PubMed=10506124;
RA Chairoungdua A., Segawa H., Kim J.Y., Miyamoto K.-I., Haga H.,
RA Fukui Y., Mizoguchi K.-I., Ito H., Takeda E., Endou H., Kanai Y.,
RT *Identification of an amino acid transporter associated with the

cystinuria-related type II membrane glycoprotein.";

J. Biol. Chem. 274:28845-28848(1999).

- !- FUNCTION: INVOLVED IN THE HIGH-AFFINITY, SODIUM-INDEPENDENT TRANSPORT OF CYSTINE AND NEUTRAL AND DIBASIC AMINO ACIDS (SYSTEM B(0,+)-LIKE ACTIVITY). THOUGHT TO BE RESPONSIBLE FOR THE HIGH-AFFINITY REABSORPTION OF CYSTINE IN THE KIDNEY PROXIMAL TUBULE.

- !- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID TRANSPORT PROTEIN SLC3A1.

- !- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

- !- TISSUE SPECIFICITY: KIDNEY AND SMALL INTESTINE. IN THE KIDNEY LOCALIZED TO THE APICAL MEMBRANE OF THE PROXIMAL TUBULES.

- !- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT) SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB029559; BAA85186.1; ..

InterPro; IPR002293; AA/rel_prmeasel.

InterPro; IPR004841; Permease.

pfam; PF00324; aa_permeases; 1.

Transport; Amino-acid transport; Transmembrane.

TRANSMEM	30	50	POTENTIAL.
TRANSMEM	64	84	POTENTIAL.
TRANSMEM	100	120	POTENTIAL.
TRANSMEM	179	199	POTENTIAL.
TRANSMEM	255	275	POTENTIAL.
TRANSMEM	297	317	POTENTIAL.
TRANSMEM	349	369	POTENTIAL.
TRANSMEM	375	395	POTENTIAL.
TRANSMEM	410	430	POTENTIAL.
TRANSMEM	435	455	POTENTIAL.

SEQUENCE 487 AA; 53657 MW; EB43AA1FCC8DC3A7 CRC64;

SO

QY	706	ATCCCTGAAC	TGCGATCAAG	CTCATTTAC	AGCTGTGG	GCATAAC	TGTAGT	GATGGT	CCCTA	765	
Db			:::		:::			:	:::		
QY	146	ProProAlaVal	ValValLysLeu	LeuAlaAlaAla	AlaAlaLeu	LeuLeuIle	ThrThrVal			165	
Db											
QY	766	AATAGCATG	AGTGTG	CAGCTGG	AGCGCCG	GGATCC	AGATTTT	CTTAAC	CTTTTG	CAAGCTC	825
Db					:::						
QY	166	AsnAlaLeu	SerValArg	LeuGlySer	TyrVal	GlnAsn	ValPhe	ThrAla	AlaLys	Leu	185
Db											
QY	826	ACAGCAAT	TCTGATA	ATTATAG	TCCCTT	TTTCAG	AAAGAG	ATTCA	AGTATG	CAGCTA	885
Db		:::									
QY	186	ValIleVal	AlaAlaIle	IleIleSer	GlyLeu	ValLeu	LeuAla	GlnGly	AsnVal	Lys	205
Db											
QY	886	AACTTTAA	GACGCCT	TTTTCAG	AAAGAG	ATTCA	AGTATG	CAGCTA	AGTATG	CAGCTA	945
Db											
QY	206	AsnPheGln	AsnSerPhe	GluGlySer	GlnThr	SerVal	GlySer	IleSer	IleLeu	AlaPhe	225
Db											
QY	946	TATTATG	GAATGT	ATGCAT	ATGCTG	CTGGT	TTTAC	CTCAAC	ATTG	TGTTACT	1005
Db											
QY	226	TyrAsnGly	LeuTrpAla	TyrAsp	GlyTrp	AsnGln	LeuAsn	TyrIle	ThrGlu	Leu	245
Db											
QY	1006	GAAACCC	TGAAAAA	ACCATT	CCCCCT	TGCA	ATATG	TATATC	CAATG	CGCCATT	1065
Db											
QY	246	ArgAsnPro	TyrArg	AsnLeu	ProMet	AlaIle	ValIle	GlyIle	ProLeu	ValThr	265
Db											
QY	1066	GGCTATG	CTGCACA	AAATGT	GGCCCT	TACTT	ACGACC	ATTAA	TGCTG	AGGAGCT	1125
Db											
QY	266	CysTyrIle	LeuMet	AsnIle	AlaTyr	PheThr	ValMet	ThrPro	ThrGlu	LeuGln	285
Db											
QY	1126	TCAAATG	CAGTGG	CAGT	GACCTT	TTCTG	ACGGCT	ACTG	GGAA	AAATTT	1185
Db											
QY	286	SerGlnAla	ValAlaVal	ThrPhe	GlyAsp	ArgVal	LeuTyr	ProAla	SerTrp	ValVal	305
Db											
QY	1186	CGCATCT	TTGTTG	CCCTCT	CCCTG	CTTTG	GCCTC	CAATG	AACG	GTGGTGT	1245
Db											
QY	306	ProLeuPhe	ValAlaPhe	SerThr	IleGly	AlaAla	AsnGly	ThrCys	PheThr	AlaGly	325
Db											
QY	1246	AGGTTAT	TCTATG	TGTCGT	CTCG	AGAGG	GTTC	ACCTT	CCAGA	AAATC	1305
Db											
QY	326	ArgLeuIle	TyrVal	AlaGly	ArgGlu	GlyHis	MetLeu	LysVal	LeuSer	TyrIle	345
Db											
QY	1306	GTCCGCA	AGCACAC	ACTCCT	CTAC	CAGCTG	TTATG	TTTGC	ACCCTT	TGC	1365
Db											

Db 295 MetValPheLeuAlaIleIleTyrIlePheValAsnIleAlaTyrPheAlaValValPro 314
QY 1108 GCTGAGGAGCTGCTGCTTTCAAAATGCAGTG---GCAGTGACCTTTTCTGAGCGGCTACTG 1164
Db 315 LysAspLysLeuIleSerLysLeuIleAlaAlaAspPheAspIleValPhe 334
QY 1165 ---GGAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATG 1221
Db 335 GlyGlyGlnAlaLysArgAlaAlaAlaLeuValGlyLeuSerAlaLeuGlyAsnVal 354
QY 1222 AACGGTGGTGTGCTGCTCCAGGTTATCTATGTTGGCTCTCGAGAGGGTCACTCTT 1281
Db 355 LeuSerValIlePheSerGlnGlyArgIleIleGlnGlnLeuGlyArgGluGlyValLeu 374
QY 1282 CCA-----GAAATCCTCTCCATGATTCTATGTTCCGCAAGCACACTCCTCTA 1326
Db 375 ProPheSerAsnPhePheAla-SerSerLysProPheAsnSerProMetValGlyLeuPh 394
QY 1327 CCAGCTGTTATGTTTGGACCTTTTGACAATGATATGCTCTTCTCTGGAGACCTCGAC 1386
Db 394 eGlnHisPheIleValCysThr--ValThrIleLeuAlaProProProGlyAspAlaTyr 413
QY 1387 AGTCTTTTG---AATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCT 1443
Db 414 LeuLeuValGlnAsnLeuIleSerTyrProMetAsnIleIleAsnPheAlaIleSerAla 433
QY 1444 GGGCTGATTATCTT-----CGATACAAATGCCAGATATGCATCGTCTCTTC 1491
Db 434 GlyLeuLeuTyrIleTyrTrpGlnArgArgGlnGlyLysIleGluTrpAsnProIle 453
QY 1492 AAGTGCCCACTGTTCCATCCAGCTTTGTTTCTTCTTCCATGCTCTTCTGCTGCTT 1551
Db 454 LysAlaGlyValPheValThrGlyPhePheThrLeuSerAsnLeuTyrLeuIleAla 473
QY 1552 -----TCCCTCTATTCGGAC---CCATTT----- 1572
Db 474 ProTyrValProProSerAsnGlyGluSerValTyrSerSerMetProTyrTrpIleHis 493
QY 1573 -----AGTACAGGGATTGGCTTCGTCATCACTCTGAGTGGAGTCCCTGCGTATTAT 1623
Db 494 CysValIleAlaTyrGlyIlePhePhe-----PheGlyGlyVal-----TyrTyr 508
QY 1624 CTCTTTATTATATGGGACAAG---AAACCCAGGTGG-----TTAGATAATATGTCAGAG 1674
Db 509 -----ValValTrpAlaGlnLeuLeuProArgTrpGlyHisTyrLysLeuValSerLys 526
QY 1675 AAAATA 1680
Db 527 AspVal 528

RESULT 14
MUP3_YEAST
ID MUP3_YEAST STANDARD; PRT; 546 AA.
AC P38734;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Low-affinity methionine permease.
GN MUP3 OR YHL036W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,

RA Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
RA Vaudin M.,
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RL Science 265:2077-2082(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=97049127; PubMed=8893857;
RA Isnard A.D., Thomas D., Surdin-Kerjan Y.,
RT "The study of methionine uptake in Saccharomyces cerevisiae reveals a
RT new family of amino acid permeases.";
RL J. Mol. Biol. 262:473-484(1996).
CC -!- FUNCTION: VERY LOW AFFINITY PERMEASE FOR METHIONINE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO YEAST HIGH AFFINITY METHIONINE PERMEASE (MUP1).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11583; AAB65048.1; -.
DR PIR; S48932; S48932.
DR SGD; S0001028; MUP3.
DR InterPro; IPR002293; AA/rel_prmeasel.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 320 340 POTENTIAL.
FT TRANSMEM 347 367 POTENTIAL.
FT TRANSMEM 394 414 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 460 480 POTENTIAL.
FT TRANSMEM 495 515 POTENTIAL.
SQ SEQUENCE 546 AA; 60619 MW; 9C9767A181FDDCAE CRC64;

Alignment Scores:

Pred. No.: 2.8e-19 Length: 546
Score: 343.50 Matches: 138
Percent Similarity: 42.96% Conservative: 100
Best Local Similarity: 24.91% Mismatches: 213
Query Match: 8.34% Indels: 103
DB: 1 Gaps: 25

US-09-667-170A-440 (1-2239) x MUP3_YEAST (1-546)

QY 247 CCTTCCCTGGCAAC-----AAGGAGCCACCTGGGCGAGGAGCGCTTT----- 288
Db 3 ProLeuLeuPheAsnSerGlyLysAlaAsnProSerGlnAspValPheIleAspValGlu 22
QY 289 -----TCAGGAAGAGACGCTTTTCAGGAAGAGAC 318
Db 23 ValGlyAspIleThrThrLysTyrGlySerThrAsnThrGlySerPheSerMetAsp 42
QY 319 GCCTTTTCAGGAAGAGAGAG---AAAGTCAGCTGAAGAGGAAGATCACTTTACTGAGGGA 375
Db 43 ThrValGluAlaGlnAlaIleLysAlaGluThrAlaArgPheMetGluValProGlnGly 62
QY 376 -----GTCTCCATTATCATTTGGCACCACCATCATTTGGAGCAGGA 411
Db 63 ArgHisLeuGlyValPheSerThrValValLeuPheValSerArgIleMetGlySerGly 82
QY 412 ATCTTTCATCTCTCTAAGGGCGTCTCCAGAACACCGGCGAGCGTGGCATGTCTCTG--- 468
Db 83 IlePheAlaValProSerValIleLeuLeuAsnThrGly-----GlyAsnLysLeuIle 100

QY	469	-----ACCATCTGGACGGTGTGGGGTCTCTGTCACTATATTTGGAGCTTTGTCTTATGCT	522
Db	101	TyrPheAlaIleTrpValPheSerAlaAlaIleAlaPheAlaGlyLeuTyrLeuPheLeu	120
QY	523	GAATTGGGAACAACACTATAAGAAATCTGGAGGTCATTACACATATATTTTGGAAAGTCTTT	582
Db	121	GluPheGlySerTrpIleProLysSerGlyGlyArgLysAsnPheLeuGluArgSerPhe	140
QY	583	GGTCCATTACCAAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAATACCCCTGCAGCT	642
Db	141	GluArgProArgLeuLeuIleSerValValPheSerCysTyrSerValLeuThrGlyTyr	160
QY	643	ACTGCTGTGATATCCCTGGCAATTTGGACGCTACATTCTGGAAACCATTTTATTCAATGT	702
Db	161	AlaLeuThrGlySerIleValPheGlyLysTyrValLeuSerAlaPheGlyVal-----	178
QY	703	GAATCCCTGAACTTGCATCAAGCTCATACAGCTGTGGGCATAACTGTAGTGATGGTC	762
Db	179	---ThrAspAspSerTrpSerLysTyrValSerIleSerPheIleIlePheAlaValLeu	197
QY	763	CTAAATAGCATGAGTGTCAAGCTGGAGCGCCGGATCCAGATTCTTTAACCTTTTGC AAG	822
Db	198	IleHisGlyValSerValArgHisGlyValPheIleGlnAsnAlaLeuGlyGlyLeuLys	217
QY	823	CTCACAGCAATTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATT-----	870
Db	218	LeuIleMetIleValLeuMetCysPheAlaGlyLeuTyrThrLeuPhePheTyrLysSer	237
QY	871	AAAGGTCAA-----ACGCAGAACCTTTAAAGACGCCTTTTTCAGGA	909
Db	238	ThrGlyGlnValAlaTrpAspLeuProValThrGlnValGluLysAspSerLeu-----	255
QY	910	ACAGATTCAAGTATTACCGCGGTGGCCACTGGCTTTTATTATGGAATGTATGCATATGCT	969
Db	256	-----LeuSerValSerSerIleAlaThrAlaPheIleSerSerPhePheCysPheSer	273
QY	970	GGCTGGTTTACCTCACTTTGTTACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTC	1029
Db	274	GlyTrpAspThrValHisThrValThrSerGluIleLysAsnProValLysThrLeuLys	293
QY	1030	CTTGCAATATGTATATCCATGGCCATTGTCCACCATGGCTATGTGCTGACAAATGTGGCC	1089
Db	294	ValSerGlyProLeuSerLeuIleIleCysPheValCysTyrThrMetMetAsnValAla	313
QY	1090	TACTTTACGACCATTAATGCTGAGGAGCTGCTG-----CTTTCAAATGCAGTG	1137
Db	314	TyrLeuLysValLeuThrTyrGluGluIleValSerAlaGlyProLeuValGlySerVal	333
QY	1138	GCAGTGACC-----TTTCTGACGGCTACTGGGAAATTTCTCATTAGCAGTCCGATC	1191
Db	334	LeuPheThrLysLeuPheGlyProArgValGlyGlyLysPhe-----IleAlaPhe	350
QY	1192	TTTGTGGCCCTCTCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGTCTGCCAGGTTA	1251
Db	351	SerIleAlaIleSerAlaAlaSerAsnIleLeuValvalIleTyrSerIleSerArgVal	370
QY	1252	TTCTATGTTGCGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCCGC	1311
Db	371	AsnGlnGluIlePheLysGluGlyTyrLeuPro-----PheSerIleHisMetSer	387
QY	1312	AAGCAC-----ACTCCTCTACCAAGCTGTTATTGTTTTCACCCCTTTTGACAATG	1359
Db	388	LysAsnTrpProPheAspAlaProLeuProSerIleSerLeuCysGlyPheIleThrIle	407
QY	1360	-----ATAATGCTCTCTCTGGAGACCTCGACAGT-----CTTTTGAATTTCTC	1404
Db	408	AlaTrpIleLeuIleLeuProLysGluGlyGluSerPheAsnTyrLeuValSerMetAsp	427
QY	1405	AGTTTGGCAGGTGCTTTTATTGGGCTGGCAGTTGTCTGGGCTGATTTATCTTCGATAC	1464
Db	428	GlyTyrGlyAsnGlnPhePheLeuLeuValAlaIleGlyLeuPheIleTrpArgPhe	447

FT TRANSMEM 355 375 POTENTIAL.
 FT TRANSMEM 389 410 POTENTIAL.
 FT TRANSMEM 417 435 POTENTIAL.
 SQ SEQUENCE 445 AA; 47576 MW; B98DF5944A5E1FEF CRC64;

Alignment Scores:
 Pred. No.: 1.6e-17 Length: 445
 Score: 321.00 Matches: 113
 Percent Similarity: 44.62% Conservative: 90
 Best Local Similarity: 24.84% Mismatches: 226
 Query Match: 7.79% Indels: 26
 DB: 1 Gaps: 10

US-09-667-170A-440 (1-2239) x YHFM_ECOLI (1-445)

QY	343	CAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATCAT	402
Db	5	GluLeuGlnArgLysLeuGlyPheTrpAlaValLeuAlaIleAlaValGlyThrThrVal	24
QY	403	GGAGCAGGAATCTTCACTCTCTAAGGGCGTGCTCCAGAACACGGGCAGCGTGGGCATG	462
Db	25	GlySerGlyIlePheValSerValGlyGluValAlaLysAlaAlaGlyThrProTrpLeu	44
QY	463	TCTCTGACCATCTGAGCGGTGTGTGGGTCTCTCACTATTITGGAGCTTGTCTTATGCT	522
Db	45	ThrValLeuAlaPheValIleGlyGlyLeuIleValIleProGlnMetCysValTyrAla	64
QY	523	GAATTGGGAACAACATAAAGAAATCTGGAGGTCACTTACACATATATTTTGAAGTCTTT	582
Db	65	GluLeuSerThrAlaTyrProGluAsnGlyAlaAspTyrValTyrLeuLysAsnAlaGly	84
QY	583	GGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAACGCCCTGCAGCT	642
Db	85	SerArgProLeuAlaPheLeuSerGlyTrpAlaSerPheTrpAlaAsnAspAlaProSer	104
QY	643	ACTGCTGTGATACCCCTGGCATTGTGGACGCTACATTTCTGGAACCATTTTTTATTCAATGT	702
Db	105	LeuSerIleMetAlaLeuAla-----IleValSerAsnLeuGlyPheLeuThr	120
QY	703	GAAATCCCTGAACCTTGGGATCAAGCTCATACAGCTGTGGGCATAACTGTAGTGATGGTC	762
Db	121	ProIleAspProLeuLeuGlyLysPheIleAlaAlaGlyLeuIleIleAlaPheMetLeu	140
QY	763	CTAATAGCATGAGTGTGAGTCAGCGCCCGGATCCAGATTTTCTTAACCTTTTTCGAAG	822
Db	141	LeuHisLeuArgSerValGluGlyGlyAlaAlaPheGlnThrLeuIleThrIleAlaLys	160
QY	823	CTCAGACCAATTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAAACG	882
Db	161	IleIleProPheThrIleValIleGlyLeuGlyIlePheTrpPhe-----LysAla	177
QY	883	CAGAACTTTAAAGACGCCTTTTCA-----GGAAGAGATTCAAGTATTACGCGGTG	933
Db	178	GluAsnPheAlaAlaProThrThrThrAlaIleGlyAlaThrGlySerPheMetAlaLeu	197
QY	934	CCACTGGCTTTTATATGAATGTATGCATATATGCTGGCTGGTTTACCTCAACTTTGTT	993
Db	198	LeuAlaGlyIleSerAlaThrSerTrpSerTyrThrGlyMetAlaSerIleCysTyrMet	217
QY	994	ACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTTCCCTTGGCAATATGTATATCCATGGCC	1053
Db	218	ThrGlyGluIleLysAsnProGlyLysThrMetProArgAlaLeuIleGlySerCysLeu	237
QY	1054	ATTGTACCAATTGGCTAT-----GTGCTGACAAATGTGGCCTACTTTACG	1098
Db	238	LeuValLeuValLeuTyrThrLeuLeuAlaLeuValIleSerGlyLeuMetProPheAsp	257
QY	1099	ACCATTAAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTCACTTTTCTGAGCGG	1158
Db	258	LysLeuAlaAsnSerGluThrProIleSerAsp-----AlaLeuThrTrpIleProAla	275
QY	1159	CTACTGGGAAATTTCTCATTTAGCAGTTCCGATCTTTTGTGGCCTCTCTCTGCTTTGGCTCC	1218

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 15:45:20 ; Search time 96.5 Seconds
(without alignments)
9561.440 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=cgn2_1/USPTO_spool/US09667170/runat_04042003_090917_20296/app_query.fasta_1.2375
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPEXT=0
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_122_@runat_04042003_090917_20296 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archheap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	58.8	494	4 Q9BYH2	Q9byh2 homo sapien

2	1221	29.6	505	6	Q9TU26
3	1216	29.5	507	13	O57439
4	1213	29.4	507	13	O13020
5	1202	29.2	512	11	O8R0X8
6	1189.5	28.9	507	4	Q96QB2
7	1156.5	28.1	515	4	Q92536
8	1132.5	27.5	510	11	Q9QWS1
9	1131.5	27.5	510	11	Q921K8
10	1131.5	27.5	512	11	Q9R0S5
11	1131.5	27.5	512	11	Q9QZ66
12	1113.5	27.0	505	5	Q9V9Y0
13	1106	26.8	499	5	Q9Y1A7
14	1095	26.6	535	6	Q9N1Q4
15	1091	26.5	517	5	Q9VB75
16	1081.5	26.2	487	6	Q9N1R6
17	1080.5	26.2	500	5	Q9VKC2
18	1058.5	25.7	493	5	Q19834
19	1051.5	25.5	530	11	Q9JMH8
20	1039	25.2	523	4	Q9NS82
21	1026	24.9	493	5	O17395
22	999.5	24.3	503	5	Q26594
23	938.5	22.8	445	5	Q9VY26
24	910	22.1	464	5	Q19151
25	717	17.4	622	3	Q9HED4
26	715	17.4	550	5	Q960A3
27	715	17.4	566	5	O01831
28	708.5	17.2	361	11	Q9CW25
29	686	16.7	478	11	Q91WN3
30	672.5	16.3	662	5	O44832
31	665.5	16.2	541	5	Q9NA91
32	623	15.1	521	5	Q94197
33	622.5	15.1	562	5	Q22448
34	597.5	14.5	470	4	Q8TCU3
35	594.5	14.4	470	4	Q96N84
36	564	13.7	440	5	Q22397
37	521.5	12.7	438	16	O34739
38	507.5	12.3	465	11	Q8VIE6
39	502.5	12.2	465	11	Q8R2J1
40	482.5	11.7	273	11	Q9D8H4
41	479.5	11.6	440	16	Q99U51
42	463	11.2	455	16	Q8YWT1
43	430	10.4	92	4	Q9P1X0
44	426.5	10.4	203	4	Q96GR5
45	412.5	10.0	180	7	Q9GIP4

ALIGNMENTS

RESULT 1
Q9BYH2
ID Q9BYH2
AC Q9BYH2; PRELIMINARY; PRT; 494 AA.
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Cystine/glutamate exchanger.
GN HXCT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.Y., Chairoungdua A., Cha S.H., Segawa H., Matsuo H., Kim D.K.,
RA Endou H., Kanai Y.;
RT "Human cystine/glutamate exchanger: cDNA cloning and upregulation by
RT oxidative stress in glioma cells."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB040875; BAB40574.1;
DR InterPro; IPR002293; AA/rel_primeasel.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.

DR TIGRFAMS; TIGR00911; 2A0308; 1.
SQ SEQUENCE 494 AA; 54423 MW; BEC57A7EA4DDB0B1 CRC64;

Alignment Scores:
Pred. No.: 8,38e-205 Length: 494
Score: 2424.00 Matches: 478
Percent Similarity: 96.37% Conservative: 0
Best Local Similarity: 96.37% Mismatches: 0
Query Match: 58.83% Indels: 18
DB: 4 Gaps: 1

US-09-667-170A-440 (1-2239) x Q9BYH2 (1-494)

QY 184 AAGCCTGTTGTCCACCATCTCCAAAGGAGGTTACCTGCAGGGAATGTTAACGGGAGG 243
DB 4 LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlyAsnValAsnGlyArg 23
QY 244 CTGCTTCCCTGGGCAACAAGGAGCCACCTGGGCGAGGACGCCCTTTTCAGGAAGAGACGCC 303
DB 24 LeuProSerLeuGlyAsnLysGluProProGlyGln----- 35
QY 304 TTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
DB 36 -----GluLysValGlnLeuLysArgLysValThr 45
QY 364 TTACTGAGGGAGTCTCCATTATCATTTGTCACCATCATTTGGAGCAGGAAATCTTCTCTCT 423
DB 46 LeuLeuArgGlyValSerIleIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 65
QY 424 CCTAAGGGCTGCTCCAGAACACGGGACGGTGGGCGATGCTCTGACCACTCTGGACGGTG 483
DB 66 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal 85
QY 484 TGTGGGCTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAG 543
DB 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrIleLys 105
QY 544 AAATCTGGAGGTCAATTACACATATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTGTGTA 603
DB 106 LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal 125
QY 604 CGAGTCTGGGTGGAACCTCCTCATATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCA 663
DB 126 ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla 145
QY 664 TTTGGACGCTACATCTCTGGAACCATTTTATTTCAATGTGAATCCCTGAACCTTGGCATC 723
DB 146 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 165
QY 724 AAGCTCATTAACAGCTGTGGGCATAAAGTGTAGTGTGATGCTGTGATGATGATGATGATGAT 783
DB 166 LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 185
QY 784 TGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGCAGGCTCAGCAATCTGTGATAATT 843
DB 186 TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIleIle 205
QY 844 ATAGTCCCTGGAGTTATGAGCTAATTAAGGTCAAACGCAGAACTTAAAGACGCCCTTT 903
DB 206 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPhelLysAspAlaPhe 225
QY 904 TCAGGAAGAGATTCAAGTATTACGGGTTTGGCACTGGCTTTTATTTATGGAATGTATGCA 963
DB 226 SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 245
QY 964 TATGCTGGCTGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAGAACCTGAAAAAAC 1023
DB 246 TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 265
QY 1024 ATCCCTTGCATATGATATCCATGGCCATTGTCCACCATTTGGCTATGCTGACAAAT 1083
DB 266 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 285

QY 1084 GTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGACAGTGGCAGTG 1143
DB 286 ValAlaTyrPheThrThrIleAsnAlaGluGluLeuLeuLeuSerAsnAlaValAlaVal 305
QY 1144 ACCTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTC 1203
DB 306 ThrPheSerGluArgLeuLeuGlyAsnPheserLeuAlaValProIlePheValAlaLeu 325
QY 1204 TCCTGCTTTGGCTCCATGAACGGTGGTGTGTTCTGCTGTCTCCAGGTTATTTCTATGTTGG 1263
DB 326 SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla 345
QY 1264 TCTCGAGAGGGTCACCTTCCAGAAATCTCTCCATGATTCATGTCGCCAGCACACTCCT 1323
DB 346 SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro 365
QY 1324 CTACCAGCTGTTATTTGTTGTCACCTTTTGACAAATGATAATGCTCTCTCTGGAGACCTC 1383
DB 366 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 385
QY 1384 GACAGTCTTTTGAATTTCTCAGTTTTCAGGTTGCGAGTGGCTTTTATTGGGCTGGCAGTTGCT 1443
DB 386 AspSerLeuLeuAsnPheserLeuSerPheAlaArgTrpLeuPheIleGlyLeuAlaValAla 405
QY 1444 GGGCTGATTTATCTTCGATACAAATGCCAGATATGATGATGCTCTTCAAGGTGCCACTG 1503
DB 406 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 425
QY 1504 TTCATCCAGCTTTGTTTTCCTTTCATGCTCTCTTTCATGTTGGTGGCTTCCCTCTATTTCG 1563
DB 426 PheIleProAlaLeuPheSerPheTyrCysLeuPheMetValAlaLeuSerLeuTyrSer 445
QY 1564 GACCCATTTAGTACAGGGATTGGCTTCATGCTCCTCTGCTGAGTGGAGTCCCTGCTATAT 1623
DB 446 AspPropheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr 465
QY 1624 CTCTTTATATATGGACAAAGAACCCAGTGGTTTAGAATAATGTCA 1671
DB 466 LeuPheIleIleTrpAspLysLysProArgTrpPheArgIleMetSer 481

RESULT 2

Q9TU26 PRELIMINARY; PRT; 505 AA.
AC Q9TU26;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Blood-brain barrier large neutral amino acid transporter.
GN LAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN MICROVASCULATURE;
RX MEDLINE=99449811; PubMed=10518579;
RA Boado R.J., Li J.Y., Nagaya M., Zhang C., Pardridge W.M.;
RT "Selective expression of the large neutral amino acid transporter at
the blood-brain barrier.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:12079-12084(1999).
DR EMBL; AF174615; AAF03694.1;
DR InterPro; IPR002293; AA/rel_primeasel.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMS; TIGR00911; 2A0308; 1.
SQ SEQUENCE 505 AA; 55110 MW; D0E6ECC11B8594E1 CRC64;

Alignment Scores: 8.06e-99 Length: 505
Pred. No.: 1221.00 Matches: 234
Score:

Percent Similarity: 67.70%
Best Local Similarity: 48.45%
Query Match: 29.64%
DB: 6
Conservative: 93
Mismatch: 153
Indels: 3
Gaps: 1

US-09-667-170A-440 (1-2239) x Q9TU26 (1-505)

QY 282 CGCCTTTTCAGGAAGAGAGAGCGCTTTTTCAGGAAGAGAGCGCTTTTTCAGGAAGAGAGAGAGT 341
Db 23 ArgGlnAlaArgGluArgMetLeu-AlaAlaArgGlyAlaAspGlyAlaGluGlyGluVa 42
QY 342 GCAGCTGAAGAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCAT 401
Db 42 lThrLeuGlnArgAsnIleThrLeuLeuAsnGlyValAlaIleIleValGlyThrIleI 62
QY 402 TGGAGCAGGAATCTTCATCTCTCTAAGGGCGTGTCTCCAGAACACGGCGAGCGTGGCAT 461
Db 62 eGlySerGlyIlePheValThrProThrGlyValLeuLysGluAlaGlySerProGlyLe 82
QY 462 GTCTCTGACCATCTGGACGGTGTGTGGGCTCCTGTCTCACTATTGGAGCTTTTGTCTTATGC 521
Db 82 uAlaLeuValValTrpAlaValCysGlyValPheSerIleValGlyAlaLeuCysTyrAl 102
QY 522 TGAATTGGGAACAACTATAAAGAAATCTGGAGGTCACTTACACATATATTTTGGAGTCTT 581
Db 102 aGluLeuGlyThrIleThrLysSerGlyGlyAspTyrAlaTyrMetLeuGluValTy 122
QY 582 TGGTCCATTACCAAGCTTTTGTACGAGTCTGGTGGAGTCTCTCAATATACGCCCTGCAGC 641
Db 122 rGlySerLeuProAlaPheLeuLysLeuTrpIleGluLeuIleIleArgProSer 142
QY 642 TACTGTCTGTATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTTATTCAATG 701
Db 142 rGlnTyrIleValAlaLeuValPheAlaThrTyrLeuLeuLysProIlePheProThrCy 162
QY 702 TGAATCCCTGAACTTGGCATCTGACATCAAGCTCATTTACAGTGTGGGCATTAACCTGATGGT 761
Db 162 sProValProGluGluAlaAlaLysLeuValAlaCysLeuCysValLeuLeuThrAl 182
QY 762 CCTAAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTTCTTAACTTTTGCAA 821
Db 182 aValAsnCysTyrSerValLysAlaAlaThrArgValGlnAspAlaPheAlaAlaAlaLy 202
QY 822 GCTCACAGCAATCTGTATATATAGTCCCTGGAGTATGAGCTAATTAAGGTCAAAC 881
Db 202 sLeuLeuAlaLeuAlaLeuIleLeuLeuGlyPheIleGlnIleGlyLysGlyAspVa 222
QY 882 GCAGAACTTT-----AAGACGCTTTTTCAGGAAGAGATTTCAAGTATTACGGGTTGCC 935
Db 222 lAlaAsnLeuAspProLysSerSerPheGluGlyThrLysLeuAspValGlyAsnIleVa 242
QY 936 ACTGGCTTTTATTATGGAATGATGATATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 995
Db 242 lLeuAlaLeuTyrSerGlyLeuPheAlaTyrGlyGlyTrpAsnTyrLeuAsnPheValTh 262
QY 996 TGAAGAAGTAGAAACCTGAAACCAACCAATCCCTTGGCAATATGATATATCCATGGCCAT 1055
Db 262 rGluGluMetIleAsnProTyrArgAsnLeuProLeuAlaIleIleIleSerLeuProIl 282
QY 1056 TGTCACCATTTGGCTATGCTGACAAATGFGCCCTACTTTTACCAACATTAATGCTGAGGA 1115
Db 282 eValThrLeuValTyrValLeuThrAsnLeuAlaTyrPheThrThrLeuThrProGluGl 302
QY 1116 GCTGCTGCTTTCAATATGAGTGGCAGTGGACCTTTTCTGAGCGGCTACTGGGAAATTTCTC 1175
Db 302 nMetLeuThrSerGluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSe 322
QY 1176 ATTAGCAGTTCGGATCTTTTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1235
Db 322 rTrpIleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPh 342
QY 1236 TGCTGTCTCCAGTTATTCTATGTTGGTCTCTCGAGAGGGTCACTTTCAGAAATCTCTC 1295

Db 342 eThrSerSerArgLeuPhePheValGlyAlaArgGluGlyHisLeuProSerIleLeuSe 362
QY 1296 CATGATTTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTTCACCCCTTTCAC 1355
Db 362 rMetIleHisProArgLeuLeuThrProValProSerLeuValPheThrCysIleMetTh 382
QY 1356 AATGATAATGCTCTCTCTGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCGCCAG 1415
Db 382 rLeuLeuTyrAlaPheSerLysAspIlePheSerValIleAsnPhePheSerPhePheAs 402
QY 1416 GTGGCTTTTATTGGCTGGCAGTGTGGCTGCTGATTTATCTTCGATACAAATGCCCGAGA 1475
Db 402 nTrpLeuCysValAlaLeuAlaIleAlaGlyMetLeuTrpLeuArgTyrGlnLysProGl 422
QY 1476 TATGCATCGTCCTTTCAAGGTGCCACTGTTTCATCCCGAGCTTTGTTTCTTCATGCCT 1535
Db 422 uLeuGluArgProIleLysValHisLeuAlaLeuProValPhePheIleLeuAlaCysLe 442
QY 1536 CTTTCATGGTGGCTTTTCCCTCTCTATTCGGACCCCATTTAGTACAGGGATTGGCTTCGTAT 1595
Db 442 uPheLeuIleAlaValSerPheTrpLysThrProValGluCysGlyIleGlyPheThrIl 462
QY 1596 CACTCTGACTGGAGTCCCTGCTATTTATCTCTTTATATATATATATATATATATATATAT 1655
Db 462 eIleLeuSerGlyLeuProValTyrPheLeuGlyValTrpTrpArgAspLysProLysTr 482
QY 1656 GTTTAGAATAATGTGAGAGAAATAACCCAGAACATTAACAATAATACTGGAAGTTGTACC 1715
Db 482 pLeuLeuGlnSerIlePheSerThrThrValLeuCysGlnLysLeuMetHisValValPr 502
QY 1716 AGAAGAA 1722
Db 502 oGlnGlu 504

RESULT 3

O57439
ID O57439 PRELIMINARY; PRT; 507 AA.
AC O57439;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE L amino acid transporter-1 LAT-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99047611; PubMed=9829974;
RA Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J.,
RA Shi Y.B., Zorzano A., Palacin M.;
RT "Identification and characterization of a membrane protein (y+L amino
RT acid transporter-1) that associates with 4F2hc to encode the amino
RT acid transport activity y+L. A candidate gene for lysinuric protein
RT intolerance.";
RL J. Biol. Chem. 273:32437-32445(1998).
DR EMBL; AF019906; AAB93541.1; -.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00911; 2A0308; 1.
SQ SEQUENCE 507 AA; 55595 MW; EA0E9782B3313B82 CRC64;

Alignment Scores:
Pred. No.: 2,23e-98 Length: 507
Score: 1216.00 Matches: 230
Percent Similarity: 65.99% Conservative: 94
Best Local Similarity: 46.84% Mismatches: 149
Query Match: 29.51% Indels: 18
DB: 13 Gaps: 2

US-09-667-170A-440 (1-2239) x 057439 (1-507)	
QY	256 GGCAACAAGGAGCCACCTGGGAGGAGCGCTTTTTCAGGAAGAGAGCGCTTTTTCAGGAAGA 315
Db	:
QY	32 GlyAsnAlaGluProLysSerGlyAspGlyAlaAla----- 43
Db	:
QY	316 GACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTTACTGAGGGGA 375
Db	:
QY	44 -----ValGluLeuGlnArgThrIleThrLeuValAsnGly 55
Db	:
QY	376 GTCCTCATATATGTCACCATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGCTG 435
Db	:
QY	56 ValAlaIleValGlyThrIleIleGlySerGlyIlePheValThrProThrGlyVal 75
Db	:
QY	436 CTCGAGAACACGGGAGCGTGGGCGATGCTCTGACCATCTGGACGCTGTGTGGGCTCCTG 495
Db	:
QY	76 LeuArgGluAlaGlySerProGlyLeuSerLeuLeuValTrpAlaValCysGlyLeuPhe 95
Db	:
QY	496 TCACATATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACACTATAAGAAATCTGGAGGT 555
Db	:
QY	96 SerIleValGlyAlaLeuLeuCysTyrAlaGluLeuGlyThrThrIleSerLysSerGlyGly 115
Db	:
QY	556 CATTACACATATATTTTGGAGTCTTTTGGTCCATTACACAGCTTTTGTGACGAGTCTGGGTG 615
Db	:
QY	116 AspTyrAlaTyrValLeuGluValTyrGlyAlaLeuProAlaPheLeuLysLeuTrpVal 135
Db	:
QY	616 GAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTAC 675
Db	:
QY	136 GluLeuLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPheAlaThrTyr 155
Db	:
QY	676 ATCTGTGGAACCATTTTATTCAATGTGAAATCCCTGAACTTGGCATCAACGCTCATTTACA 735
Db	:
QY	156 LeuLeuLysProValPheProThrCysProValProAspAlaAlaLysIleValAla 175
Db	:
QY	736 GCTGTGGGCATAACTGTAGTATGGTCTCTAAATAGCATGAGTGTGCTGAGGAGCGCCGG 795
Db	:
QY	176 CysLeuCysIleLeuLeuLeuThrAlaIleAsnCysTyrSerValLysAlaAlaThrArg 195
Db	:
QY	796 ATCCAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATAATTTATAGTCCCTGGA 855
Db	:
QY	196 ValGlnAspArgPheAlaAlaLysLeuLeuAlaLeuLeuIleIleLeuGly 215
Db	:
QY	856 GTTATGCAGCTAATAAGGTCAAACGCGAGAACTTTTAAAGAC-----GCCTTTTCAGGA 909
Db	:
QY	216 PheValGlnLeuGlyLysGlyGlyValGluAspLeuLysProGluArgSerPheGluGly 235
Db	:
QY	910 AGAGATTCAAGTATTACGCGGTGTGCCACTGGCTTTTATTATGGAATGTATGCTATATGCT 969
Db	:
QY	236 ThrSerThrAsnValGlyGlnTrpValLeuAlaLeuTyrSerGlyLeuPheAlaTyrGly 255
Db	:
QY	970 GGCTGGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAAACCTGAAAACCACTCC 1029
Db	:
QY	256 GlyTrpAsnTyrLeuAsnPheValValGluGluMetIleGluProTyrLysAsnLeuPro 275
Db	:
QY	1030 CTTGCAATATGTATATCCATGGCCATTTGTCACCATTTGGCTATGCTGACAAAATGTGGCC 1089
Db	:
QY	276 ArgAlaIleIleSerMetProIleValThrLeuValTyrValLeuThrAsnLeuAla 295
Db	:
QY	1090 TACTTTACGACCATTAATCTGAGGAGCTGCTGCTTTCAAATGCGAGTGGCAGTGACCTTT 1149
Db	:
QY	296 TyrPheThrThrLeuThrProGluGlnMetLeuAsnSerGluAlaValAlaValAspPhe 315
Db	:
QY	1150 TCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCCGATCTTTGTCCTCTCTCCTGC 1209
Db	:
QY	316 GlyAsnTyrHisLeuGlyValMetAlaTrpIleIleProValPheValGlyLeuSerCys 335
Db	:
QY	1210 TTTGGCTCCATGAACGGTGGTGTGTTTGGCTGTCTCCAGGTTATTTCTATGTTGCGTCTCGA 1269
Db	:
QY	336 PheGlySerValAsnGlySerLeuPheThrSerArgLeuPhePheValGlyAlaArg 355
Db	:
QY	1270 GAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCACACTCTCTACCA 1329
Db	:
QY	356 GluGlyHisLeuProSerLeuLeuAlaMetIleHisProArgLeuLeuThrProMetPro 375
Db	:

QY	1330 GCTGTTATTGTTTTCACCCCTTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGT 1389
Db	:
QY	376 SerLeuIlePheThrCysAlaMetThrLeuLeuTyrAlaPheSerAspIlePheSer 395
Db	:
QY	1390 CTTTGAATTTCTCAGTTTTCGCCAGGTGGCTTTTATTATTTGGCTGGCAGTGTGGCTG 1449
Db	:
QY	396 ValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIleIleGlyMet 415
Db	:
QY	1450 ATTTATCTTCGATACAAATGCCAGATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATC 1509
Db	:
QY	416 MetTrpLeuArgTyrLysLysProGluLeuGluArgProIleLysValAsnIleLeuLeu 435
Db	:
QY	1510 CCAGCTTTGTTTCTCTTCACATGCCTCTTTCATGTTGGCTTTCCTCTCTATTCGGACCCA 1569
Db	:
QY	436 ProIlePhePheIleLeuAlaCysIlePheLeuIleValSerPheTyrMetThrPro 455
Db	:
QY	1570 TTTAGTACAGGATTTGGCTTCGTCATCACCTCTGACTGAGTCCCTCGGTATATCTCTTT 1629
Db	:
QY	456 ValGluCysGlyIleGlyPheIleIleLeuThrGlyValProValTyrPhePheGly 475
Db	:
QY	1630 ATTATATGGGACAAAGAAACCCAGGTGGTTTAGAAATATGTTCAGAGAAAATAACCAACA 1689
Db	:
QY	476 ValTrpTrpGlnAsnLysProAspTrpIleLeuHisGlyIleHisSerSerThrAlaLeu 495
Db	:
QY	1690 TTACAAATAATACTGGAAGTTGTACCAAGAA 1722
Db	:
QY	496 LeuGlnLysValMetGluAlaValProGlnGlu 506
Db	:
RESULT 4	
O13020 PRELIMINARY; PRT; 507 AA.	
ID	O13020
AC	O13020;
DT	01-JUL-1997 (TrEMBLrel. 04, Created)
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Amino acid transporter chain (AmAt-L-1c) ASUR4.
GN	ASUR4.
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=8355;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=KIDNEY;
RX	MEDLINE=97324256; PubMed=9178633;
RA	Spindler B., Mastroberardino L., Custer M., Verrey F.;
RT	"Characterization of early aldosterone-induced RNAs in A6 kidney epithelia";
RL	Pflugers Arch. 434:323-331(1997).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=KIDNEY;
RX	MEDLINE=98421678; PubMed=9751058;
RA	Mastroberardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffing J.;
RA	Shoemaker C.B., Verrey F.;
RT	"Amino acid transport by heterodimers of 4F2hc/CD98 and members of a new permease family";
RL	Nature 395:288-291(1998).
DR	EMBL; Y12716; CAA73254.1; -
DR	InterPro; IPR002293; AA/rel_prmeasel.
DR	InterPro; IPR004760; L_AA_transport.
DR	InterPro; IPR004841; Permease.
DR	Pfam; PF00324; aa_permeases; 1.
DR	TIGRFAMS; TIGR00911; 2A0308; 1.
FT	CONFLICT 42 42 G -> A (IN REF. 0).
FT	CONFLICT 199 199 A -> R (IN REF. 0).
FT	CONFLICT 209 209 F -> L (IN REF. 0).
FT	CONFLICT 391 391 N -> D (IN REF. 0).
SQ	SEQUENCE 507 AA; 55529 MW; 2A6741CE89AB721D CRC64;
Alignment Scores:	

QY	1210	TTTGGCTCCATGAACGGTGGTGTGTTTGCTGTCTCCAGGTATTCTATGTTGGCTCTCGA	1269
Db	336	PheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheValGlyAlaArg	355
QY	1270	GAGGTCACCTTCCAGAAATCCTCTCCATGATTGATGTCGCGAAGCACACTCCCTCTACCA	1329
Db	356	GluGlyHisLeuProSerLeuLeuAlaMetIleHisProArgLeuLeuThrProMetPro	375
QY	1330	GCTGTATTGTTTGCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACCTCGACAGT	1389
Db	376	SerLeuIlePheThrCysAlaMetThrLeuLeuTyrrAlaPheSerAsnAspIlePheSer	395
QY	1390	CTTTTGAATTTCCCTCAGTTTGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTG	1449
Db	396	ValIleAsnPhePheSerPhePheAsnTrpLeuCysValAlaLeuAlaIleIleGlyMet	415
QY	1450	ATTTATCTTCGATACAAATGCCACAGATATGCATCGTCTTTTCAAGGTGCCACTGTTCATC	1509
Db	416	MetTrpLeuArgTyrrLysLysProGluLeuGluArgProIleLysValAsnIleLeuLeu	435
QY	1510	CCAGCTTTGTTTTCCTTCACATGCCTCTTTCATGGTGGCCCTTTCCCTCTATTTCGGACCCA	1569
Db	436	ProIlePhePheIleLeuAlaCysIlePheLeuIleValSerPheTyrrMetThrPro	455
QY	1570	TTTAGTACAGGATTGGCTTCGCTCATCTGCTGACTGGAGTCCCTGCGTATTATCTCTTT	1629
Db	456	ValGluCysGlyIleGlyPheIleIleIleLeuThrGlyValProValTyrrPhePheGly	475
QY	1630	ATTATATGGGACAGAAACCCAGGTGGTTTACAATAATGTGAGAGAAATAACCAAGACA	1689
Db	476	ValTrpTrpGlnAsnLysProAspTrpIleLeuHisGlyIleHisSerSerThrAlaLeu	495
QY	1690	TTACAAATAATACTGGAAGTTGTACCAGAGAA	1722
Db	496	LeuGlnLysValMetGluAlaValProGlnGlu	506
RESULT 5			
Q8R0X8			
ID	Q8R0X8	PRELIMINARY;	PRT; 512 AA.
AC	Q8R0X8;		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 5.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=SALIVARY GLAND;		
RA	Strausberg R.;		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC026131; AAH26131.1; -.		
SQ	SEQUENCE 512 AA; 55872 MW; 57045EC4DD9DE1A1 CRC64;		
Alignment Scores:			
Pred. No.:	3.82e-97	Length:	512
Score:	1202.00	Matches:	229
Percent Similarity:	67.36%	Conservative:	95
Best Local Similarity:	47.61%	Mismatches:	151
Query Match:	29.17%	Indels:	6
DB:	11	Gaps:	2
US-09-667-170A-440 (1-2239) x Q8R0X8 (1-512)			
QY	298	GAGCCCTTTTCAGGAAGACAGCGCCTTTTTCAGGAAGAGAGAAAAGTGCAGCTGAAGAGGAAA	357
Db	31	GluAlaArgArgGlyAspGlyAlaAspProGluGlyGluGlyValThrLeuGlnArgAsn	50
QY	358	GTCACCTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTC	417

Db 51 IleThrLeuLeuAsnGlyValAlaIleValGlyThrIleIleGlySerGlyIlePhe 70
QY 418 ATCTCTCTAAGGCGTGTCTCCAGAACACGGGCGAGCGTGGCATGTCTCTGACCATCTGG 477
Db 71 ValThrProThrGlyValLeuLysGluAlaGlySerProGlyLeuSerLeuValValTrp 90
QY 478 ACGGTGTGGGGTCTGTCACTATTGGAGCTTTTGTCTTATGCTGAATTGGGAACAAC 537
Db 91 AlaValCysGlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThr 110
QY 538 ATAAAGAAATCTGGAGGTCAATTACACATATATTTTGGAGTCTTTGGTCCATTACCACT 597
Db 111 IleSerLysSerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAla 130
QY 598 TTTGTACGAGTCTGGTGGAACTCCCTCATAATACGCCCTGCAGTACTGCTGTGATATCC 657
Db 131 PheLeuLysLeuTrpIleGluLeuLeuIleIleArgProSerSerGlnTyrIleValAla 150
QY 658 CTGGCATTTGGACGCTACATTCTGGGAACCATTTTATTCAATGTGAATCCCTGAACCT 717
Db 151 LeuValPheAlaThrTyrLeuLeuLysProValPheProThrCysProValProGluGlu 170
QY 718 GCGATCAAGCTCATACAGCTGTGGGCATAACTGTAGTGATGGTCTCTAAATAGCATGAGT 777
Db 171 AlaAlaLysLeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSer 190
QY 778 GTCAGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAATCTG 837
Db 191 ValLysAlaAlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAla 210
QY 838 ATAATATAGTCCCTGGAGTATGACAGTAAATAA-----GGTCAAACGCAG 885
Db 211 LeuIleIleLeuLeuGlyPheIleGlnMetGlyLysAspMetGlyGlnGlyAspAlaSer 230
QY 886 AACTTT-----AAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTGCCACTG 939
Db 231 AsnLeuGlnGlnLysLeuSerPheGluGlyThrAsnLeuAspValGlyAsnIleValLeu 250
QY 940 GCTTTTATTATGGAATGTATGATATGCTGGTGGTGTACCTCAACTTGTACTGAA 999
Db 251 AlaLeuTyrSerGlyLeuPheAlaTyrGlyGlyTyrAsnTyrLeuAsnPheValThrGlu 270
QY 1000 GAAGTAGAAAACCCCTGAAAACACCATTCCTTGCATATGATATATCCATGCCATGTC 1059
Db 271 GluMetIleAsnProTyrArgAsnLeuProLeuAlaIleIleSerLeuProIleVal 290
QY 1060 ACCATTGCTGTATGTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTG 1119
Db 291 ThrLeuValTyrValLeuThrAsnLeuAlaTyrPheThrLeuSerThrAsnGlnMet 310
QY 1120 CTGCTTTCAAATGACGTGACGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTA 1179
Db 311 LeuThrSerGluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrp 330
QY 1180 GCAGTTCGATCTTTTGTGCCCCTCTCTGCTTTTGGCTCCATGAACGGTGGTGTGTTGCT 1239
Db 331 IleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThr 350
QY 1240 GTCTCCAGTTATTCTATGTTGCGTCTGAGAGGGTCACTTCCAGAAATCCTCTCCATG 1299
Db 351 SerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProSerValLeuSerMet 370
QY 1300 ATTCATGTCGCAAGCACACTCCTCTACCACTGTTATTGTTTTCACCCCTTTGACAATG 1359
Db 371 IleHisProGlnLeuLeuThrProValProSerLeuValPheThrCysIleMetThrLeu 390
QY 1360 ATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTTCTCAGTTTGGCAGGTG 1419
Db 391 MetTyrAlaPheSerArgAspIlePheSerIleIleAsnPhePheSerPheAsnTrp 410
QY 1420 CTTTATTGGCTGGCAGTGTGCTGGCTGATTTATCTTCGATACAAATGCCAGATATG 1479
Db 411 LeuCysValAlaLeuAlaIleIleGlyMetMetTrpLeuArgPheLysLysProGluLeu 430

QY 1480 CATCGTCTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTCTTCCATGCGCTCTTC 1539
Db 431 GluArgProIleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPhe 450
QY 1540 ATGGTTGCCCTTCTCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACT 1599
Db 451 LeuIleAlaValSerPheTrpLysTrpMetGluCysGlyIleGlyPheAlaIleIle 470
QY 1600 CTGACTGGAGTCCCTGCGTATTATCTCTTATATATATGGACAAACCCAGGTGTTT 1659
Db 471 LeuSerGlyLeuProValTyrPhePheGlyValTrpTrpLysAsnLysProLysTrpIle 490
QY 1660 AGAATAATGTACAGAGAAAATAACAGAACATTAACAATAATACTGGAAGTTGTACCAGAA 1719
Db 491 LeuGlnAlaIlePheSerValThrValLeuCysGlnLysLeuMetGlnValValProGln 510
QY 1720 GAA 1722
Db 511 Glu 511
RESULT 6
Q96QB2 PRELIMINARY; PRT; 507 AA.
ID Q96QB2
AC Q96QB2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Sodium-independent neutral amino acid transporter LAT1.
GN LAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nii T., Taketani Y., Tani Y., Ohkido M., Ségawa H., Kanai Y.,
RA Miyamoto K., Endou H., Takeda E.;
RT "Induction of the human sodium-independent neutral amino acid
RT transporter LAT1 (4F2 light chain) in activated T lymphocytes.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023721; BAB70708.1;
DR EMBL; AB023712; BAB70708.1; JOINED.
DR EMBL; AB023713; BAB70708.1; JOINED.
DR EMBL; AB023714; BAB70708.1; JOINED.
DR EMBL; AB023715; BAB70708.1; JOINED.
DR EMBL; AB023716; BAB70708.1; JOINED.
DR EMBL; AB023717; BAB70708.1; JOINED.
DR EMBL; AB023718; BAB70708.1; JOINED.
DR EMBL; AB023719; BAB70708.1; JOINED.
DR EMBL; AB023720; BAB70708.1; JOINED.
DR InterPro; IPR002293; AA/rel_pmeasel.
DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMS; TIGR00911; 2A0308; 1.
SQ SEQUENCE 507 AA; 55123 MW; DC6C559DA1ECE21E CRC64;

Alignment Scores:
Pred. No.: 4.81e-96 Length: 507
Score: 1189.50 Matches: 229
Percent Similarity: 65.59% Conservative: 95
Best Local Similarity: 46.36% Mismatches: 167
Query Match: 28.87% Indels: 3
DB: 4 Gaps: 2

US-09-667-170A-440 (1-2239) x Q96QB2 (1-507)

QY 247 CCTTCCCTGGCAACAAGAGCCACCTGGGACGAGCGCTTTTTCAGGAAGAGACGCTTT 306
Db 14 ProValAlaGluGluLysGluAlaArgGluLysMetLeuAlaSerLysArgAlaAsp 33
QY 307 TCAGGAAGAGACGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTA 366

Db 37 GluThrMetGlnLeuLysLysGluIleSerLeuLeuAsnGlyValSerLeuValValGly 56
QY 394 ACCATGATGGAGCAGGAATCTCATCTCTCCTAAGGCGGTGCTCCAGAACACGGGCGAGC 453
Db 57 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuValHisThrAlaSer 76
QY 454 GTGGGCGATGCTCTGACCATCTGGACGGGTGTGGGGTCTGTCTCACTATTTGGAGCTTTG 513
Db 77 TyrGlyMetSerLeuIleValTrpAlaIleGlyGlyLeuPheSerValValGlyAlaLeu 96
QY 514 TCTTATGCTGAATGGGAACAACATATAAGAAATCTGGAGGTCAATTACACATATATTTG 573
Db 97 CystyrAlaGluLeuGlyThrThrIleThrLysSerGlyAlaSerTyrAlaTyrIleLeu 116
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATACGC 633
Db 117 GluAlaPheGlyGlyPheIleAlaPheIleArgLeuTrpValSerLeuValValGlu 136
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTGGAACCATTTT 693
Db 137 ProThrGlyGlnAlaIleIleAlaIleThrPheAlaAsnTyrIleIleGlnProSerPhe 156
QY 694 ATTCAATGTGAATCCCTGAACCTGGCATCAAGCTCATTTACAGCTGTGGGCATATACTGTA 753
Db 157 ProSerCysAspProTyrLeuAlaCysArgLeuLeuAlaAlaAlaCysIleCysLeu 176
QY 754 GTGATGGTCCCTAATAGCATGAGTGTGAGTGGAGCGCCGGATCCAGATTTTCTTAACC 813
Db 177 LeuThrPheValAsnCysAlaTyrValLysTrpGlyThrArgValGlnAspThrPheThr 196
QY 814 TTTTGCAGCTCACAGCAATCTGTATATATAGTCCCTGGAGTTATGCAGCTAATATAA 873
Db 197 TyrAlaLysValValAlaLeuIleAlaIleIleValMetGlyLeuValLysLeuCysGln 216
QY 874 GGTCAACCGCAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGCGGTTG 933
Db 217 GlyHisSerGluHisPheGlnAspAlaPheGluGlySerSerTrpAspMetGlyAsnLeu 236
QY 934 CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 993
Db 237 SerLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnPheVal 256
QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTCCTCCCTTGCAATATGTATATCCATGGCC 1053
Db 257 ThrGluGluIleLysAsnProGluArgAsnLeuProLeuAlaIleGlyIleSerMetPro 276
QY 1054 ATTGTACCATTTGGCTATGCTGTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAG 1113
Db 277 IleValThrLeuIleTyrIleLeuThrAsnValAlaTyrThrValLeuAsnIleSer 296
QY 1114 GAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 297 AspValLeuSerSerAspAlaValAlaValThrPheAlaAspGlnThrPheGlyMetPhe 316
QY 1174 TCATTACGAGTCCGATCTTTGTTGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTGTG 1233
Db 317 SerTrpThrIleProIleAlaValAlaLeuSerCysPheGlyGlyLeuAsnAlaSerIle 336
QY 1234 TTTGCTGCTCCAGGTTATTCTATGTTGGCTCTCGAGAGGGTCACCTTCCAGAAATCCTC 1293
Db 337 PheAlaSerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProAspLeuLeu 356
QY 1294 TCCATGATTCATGTCGCGAAGCACACTCCTCTACCGCTGTTATTGTTTGGCCCTTTG 1353
Db 357 SerMetIleHisIleGluArgPheThrProIleProAlaLeuLeuPheAsnCysThrMet 376
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC 1413
Db 377 AlaLeuIleTyrLeuIleValGluAspValPheGlnLeuIleAsnTyrPheSerPheSer 396
QY 1414 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATACAAATGCCCA 1473
Db 397 TyrTrpPhePheValGlyLeuSerValValGlyGlnLeuTyrLeuArgTrpLysGluPro 416

QY 1474 GATATGCATCGTCTCTTCAAGGTGCCACTGTTTCAATCCAGCTTGTGTTTCTTCCATCATGC 1533
Db 417 LysArgProArgProLeuLysLeuSerValPhePheProIleValPheCysIleCysSer 436
QY 1534 CTCTTTCATGTTGCCCTTTCCTCTCTATTCGGACCCCATTTAGTACAGGATTTGGCTTCGTC 1593
Db 437 valPheLeuValIleValProLeuPheThrAspThrIleAsnSerLeuIleGlyIleGly 456
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATATCTC-----TTTATTATATGGGACAGAAA 1647
Db 457 IleAlaLeuSerGlyValProPheTyrPheMetGlyValTyrLeuProGluSerArgArg 476
QY 1648 CCCAGGTGTTAGAAATAATGTCCAGAGAAATAACACAGA--ACATTACAAATAATACTG 1704
Db 477 ProLeuPheIleArgAsnValLeuAlaAlaIleThrArgGlyThrGlnGlnLeuCysPhe 496
QY 1705 GAAGTGTACCAAGAAAGAT 1725
Db 497 CysValLeuThrGluLeuAsp 503
RESULT 8
Q9QWS1 PRELIMINARY; PRT; 510 AA.
ID Q9QWS1
AC Q9QWS1; 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Y+LAT1b.
GN SLC7A7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH/SWISS; TISSUE=HEART;
RX MEDLINE=99094891; PubMed=9878049;
RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
RT "Amino acid transport of y+L-type by heterodimers of 4F2hc/CD98 and
RT members of the glycoprotein-associated amino acid transporter
RT family.";
RL EMBO J. 18:49-57(1999).
DR EMBL; AJ130943; CAA10255.1; -.
DR MGD; MGI:1337120; SLC7a7.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
FT VARIANT 4 4 T -> S.
SQ SEQUENCE 510 AA; 55691 MW; 9F30FA7BAF126F6C CRC64;

Alignment Scores:
Pred. No.: 510 Length: 510
Score: 1132.50 Matches: 214
Percent Similarity: 68.59% Conservative: 107
Best Local Similarity: 45.73% Mismatches: 142
Query Match: 27.49% Indels: 5
DB: 11 Gaps: 3

US-09-667-170A-440 (1-2239) x Q9QWS1 (1-510)

QY 334 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGCG 393
Db 30 GluGlnValLysLeuLysLysGluIleSerLeuLeuAsnGlyValCysLeuIleValGly 49
QY 394 ACCATCATTTGGAGCAGGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 453
Db 50 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuMetTyrSerAlaSer 69
QY 454 GTGGGCGATGCTCTGACCATCTGGACGGTGTGGGGTCTCTGCTGCTGCTGCTGCTGCTGCT 513
Db 70 PheGlyLeuSerLeuValIleTrpAlaValGlyGlyIlePheSerValPheGlyAlaLeu 89

QY 514 TCATTGCTGAATTGGGAACAACACTATAAAGAAATCTGGAGGTCAATTACACATATATTTTG 573
Db 90 CysTyrAlaGluLeuGlyThrThrIleLysSerGlyAlaSerTyrAlaTyrIleLeu 109
QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAATCCTCATATAATACGC 633
Db 110 GluAlaPheGlyGlyPheLeuAlaPheIleArgLeuTrpThrSerLeuLeuIleGlu 129
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTCTGGAAACCATTTT 693
Db 130 ProThrSerGlnAlaValIleAlaIleThrPheAlaAsnTyrMetValGlnProLeuPhe 149
QY 694 ATTCAAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTCACAGCTGTGGGCATACGTGA 753
Db 150 ProSerCysGlyAlaProTyrAlaAlaGlyArgLeuLeuAlaAlaCysIleCysLeu 169
QY 754 GTGATGGTCTAAATAGCATGAGTGTACGCTGGAGCGCCCGGATCCAGATTTTCTTAACC 813
Db 170 LeuThrPheIleAsnCysAlaTyrValLysTrpGlyThrLeuValGlnAspIlePheThr 189
QY 814 TTTTGAAGCTCACAGCAATCTGTATATATATAGTCCCTGGAGTATGATATATCCATGACC 873
Db 190 TyrAlaLysValLeuAlaLeuIleAlaValIleIleAlaGlyIleValArgLeuGlyGln 209
QY 874 GGTCAAACGCAGCAACTTTAAAGACGCTTTTCAGGAAGAGATTCAGATATTACGGCGTTG 933
Db 210 GlyAlaThrAlaAsnPheGluAsnSerPheGluGlySerPheAlaMetGlyAspIle 229
QY 934 CCACTGGCTTTTATATGGAATGATGATCATATGCTGGCTGGTTTACCTCAACTTGT 993
Db 230 AlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnTyrVal 249
QY 994 ACTGAAGAAGTAGAAAACCTTGAAAAAACCATTCCTCCCTTGCATATATGATATCCATGCC 1053
Db 250 ThrGluGluIleArgAsnProGluArgAsnLeuProLeuSerIleGlyIleSerMetPro 269
QY 1054 ATTGTCACCAATGGCTATGTGTCGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 1113
Db 270 IleValThrIleIleTyrLeuLeuThrAsnValAlaTyrTyrSerValLeuAspIleLys 289
QY 1114 GAGCTGCTGCTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 290 GluIleLeuAlaSerAspAlaValAlaValThrPheAlaAspGlnIlePheGlyValPhe 309
QY 1174 TCATTAGCAGTCCGATCTTTGTGTCCTCTCTCCCTGCTTGGCTCCATGACGGTGGTGTG 1233
Db 310 AsnTrpIleProValAlaValAlaPheSerCysPheGlyGlyLeuAsnAlaSerIle 329
QY 1234 TTTGCTGCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATCCTC 1293
Db 330 ValAlaAlaSerArgLeuLeuPheValGlySerArgGlyHisLeuProAspAlaIle 349
QY 1294 TCCATGATTATGTCGCGCAGCACACTCCTCTACAGCTGTTATTGTTTGCACCCCTTG 1353
Db 350 CysMetValHisValGluArgPheThrProValProSerLeuLeuPheAsnGlyValLeu 369
QY 1354 ACAATGATAATGCTCTTCTGAGAGACCTCGACAGCTCTTTTGAATTCCTCAGTTTGGC 1413
Db 370 SerLeuValTyrLeuCysValGluAspIlePheGlnLeuIleAsnTyrTyrSerPheSer 389
QY 1414 AGGTGGCTTTTATTGGCTGGCAGTTGCTGGCTGATTATCTTCGATACAAATGCCCA 1473
Db 390 TyrTrpPheValGlyLeuSerIleValGlyGlnLeuTyrLeuArgTrpLysAspPro 409
QY 1474 GATATGCATCGTCTTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCTTCACATGC 1533
Db 410 AspArgProArgProLeuLysLeuSerLeuPhePheProIleIlePheCysLeuCysThr 429
QY 1534 CTCTTCATGGTTGCCCTTCTCCCTCTATTCCGACCCCATTTAGTACAGGATTTGGCTTCGC 1593
Db 430 IlePheLeuValAlaValProLeuTyrSerAspThrIleAsnSerLeuIleGlyIleGly 449
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTATTATA-----TGGGACAAG 1644

Db 450 IleAlaLeuSerGlyLeuPro--PheTyrPhePheIleIleArgValProGluHisLys 468
QY 1645 AAACCCAGGTGCTTTAGATAATATGTCAGAGAAATAACACGAAACATTACAAATAATA--- 1701
Db 469 ArgProLeuPheLeuArgArgIleValAlaSerIleThrArgTyrLeuGlnIleLeuCys 488
QY 1702 CTGGAAGTTGTACCAGAGAAGAT 1725
Db 489 MetSerValAlaAlaGluMetAsp 496
RESULT 9
Q9Z1K8 PRELIMINARY; PRT; 510 AA.
AC Q9Z1K8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Y+LAT1a (Solute carrier family 7 (Cationic amino acid transporter, y+ system), member 7).
GN SLC7A7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=KIDNEY;
RX MEDLINE=99094891; PubMed=9878049;
RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
RT "Amino acid transport of y+L-type by heterodimers of 4F2hc/CD98 and members of the glycoprotein-associated amino acid transporter family.";
RT EMBO J. 18:49-57(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ012754; CAA10170.1; -;
DR EMBL; BC014709; AAHL4709.1; -;
DR MGD; MGI:1337120; Slc7a7.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
FT VARIANT 4 4 S -> T.
SQ SEQUENCE 510 AA; 55677 MW; 9F30FB1B88126F6C CRC64;

Alignment Scores:

Pred. No.: 5,06e-91 Length: 510
Score: 1132.50 Matches: 214
Percent Similarity: 68.59% Conservative: 107
Best Local Similarity: 45.73% Mismatches: 142
Query Match: 27.49% Indels: 5
DB: 11 Gaps: 3

US-09-667-170A-440 (1-2239) x Q9Z1K8 (1-510)

QY 334 GAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
Db 30 GluGlnValLysLeuLysLysGluIleSerLeuLeuAsnGlyValCysLeuIleValGly 49
QY 394 ACCATCATTTGGACGAGGAATCTTCTCATCTCTCTAAGGGCGTCTCCAGAACACGGGCAGC 453
Db 50 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuMetTyrSerAlaSer 69
QY 454 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGGTCTGTGCTCACTATTGGAGCTTTG 513
Db 70 PheGlyLeuSerLeuValIleTrpAlaValGlyGlyIlePheSerValPheGlyAlaLeu 89
QY 514 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCACTTACACATATATTTG 573
Db 90 CysTyrAlaGluLeuGlyThrThrIleLysLysSerGlyAlaSerTyrAlaTyrIleLeu 109

Qy 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAACGC 633
Db 110 GluAlaPheGlyGlyPheLeuAlaPheIleArgLeuTrpThrSerLeuLeuIleGlu 129
Qy 634 CCTGCAGTACTGCTGTGATATCCCTGGCAATTTGGACGCTACATCTTGGAAACCATTTT 693
Db 130 ProThrSerGlnAlaValIleAlaIleThrPheAlaAsnTyrMetValGlnProLeuPhe 149
Qy 694 ATTCATGTGAAATCCCTGAACCTTGCATCAAGCTCAATACAGCTGTGGGCATAACTGTA 753
Db 150 ProSerCysGlyAlaProTyrAlaAlaGlyArgLeuLeuAlaAlaCysIleCysLeu 169
Qy 754 GTGATGGTCTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACC 813
Db 170 LeuThrPheIleAsnCysAlaTyrValLysTrpGlyThrLeuValGlnAspIlePheThr 189
Qy 814 TTTTGGCAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATAA 873
Db 190 TyrAlaLysValLeuAlaLeuIleAlaValIleIleAlaGlyIleValArgLeuGlyGln 209
Qy 874 GGTCAAACGAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAGTATTACCGGTTG 933
Db 210 GlyAlaThrAlaAsnPheGluAsnSerPheGluGlySerSerPheAlaMetGlyAspIle 229
Qy 934 CCAGTGGCTTTTATTATGGAATGTATGATATGCTGGCTGGTGTTCACCTCAACITTTGTT 993
Db 230 AlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnTyrVal 249
Qy 994 ACTGAAGAAGTAGAAAACCCCTGAAAACCACTTCCCTTGCATATATATATCCATGGCC 1053
Db 250 ThrGluGluIleArgAsnProGluArgAsnLeuProLeuSerIleGlyIleSerMetPro 269
Qy 1054 ATGTGTCACCATTTGGCTATGTGTGACAAATGTGGCTTCTTACGACCATTAATGCTGAG 1113
Db 270 IleValThrIleIleTyrLeuLeuThrAsnValAlaTyrTyrSerValLeuAspIleLys 289
Qy 1114 GAGCTGCTGCTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGAAATTC 1173
Db 290 GluIleLeuAlaSerAspAlaValAlaValThrPheAlaAspGlnIlePheGlyValPhe 309
Qy 1174 TCATTAGCAGTTCCGATCTTTGTTGCCCTCTCTCTGCTTGGTCCATGAACGGTGTG 1233
Db 310 AsnTrpIleIleProValAlaValAlaPheSerCysPheGlyGlyLeuAsnAlaSerIle 329
Qy 1234 TTTGCTGTCTCCAGGTTATCTATGTTGCTGCTCGAGAGGTCACCTTCCAGAAATCCTC 1293
Db 330 ValAlaAlaSerArgLeuLeuPheValGlySerArgGluGlyHisLeuProAspAlaIle 349
Qy 1294 TCCATGATTCATGTCGCGAAGCACACTCTCTACAGCTGTTTGTATTTGTCACCTTTG 1353
Db 350 CysMetValHisValGluArgPheThrProValProSerLeuLeuPheAsnGlyValLeu 369
Qy 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTTTTGAATTTCTCAGTTTGGC 1413
Db 370 SerLeuValTyrLeuCysValGluAspIlePheGlnLeuIleAsnTyrTyrSerPheSer 389
Qy 1414 AGGTGGCTTTTATTGGCTGGCAGTGTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 390 TyrTrpPhePheValGlyLeuSerIleValGlyGlnLeuTyrLeuArgTrpLysAspPro 409
Qy 1474 GATATGCATCGTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTCTCATATGC 1533
Db 410 AspArgProArgProLeuLysLeuSerLeuPhePheProIleIlePheCysLeuCysThr 429
Qy 1534 CTCCTCATGTTGGCTTCCCTCTATTCGGACCCATTTAGTACAGGGATTTGGCTTCGTC 1593
Db 430 IlePheLeuValAlaValProLeuTyrSerAspThrIleAsnSerLeuIleGlyIleGly 449
Qy 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTATTATA-----TGGGACAAAG 1644
Db 450 IleAlaLeuSerGlyLeuPro---PheTyrPhePheIleIleArgValProGluHisLys 468
Qy 1645 AAACCCAGGTGGTTTAGAATAATGTCAGAGAAAATAACAGAACATTAACAATAATAA--- 1701

Db 469 ArgProLeuPheLeuArgArgIleValAlaSerIleThrArgTyrLeuGlnIleLeuCys 488
Qy 1702 CTGGAAGTGTGTACCAGAGAAGAT 1725
Db 489 MetSerValAlaAlaGluMetAsp 496
RESULT 10
Q9R0S5 PRELIMINARY; PRT; 512 AA.
AC Q9R0S5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Y+L amino acid transporter 1.
GN RY+LAT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Kanai Y., Fukasawa Y., Segawa H., Endou H.;
RT "Characterization of a system Y+L amino acid transporter.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB020520; BAA87325.1; -
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
SQ SEQUENCE 512 AA; 55641 MW; 708905B2ACB0A824 CRC64;
Alignment Scores:
pred. No.: 6.2e-91 Length: 512
Score: 1131.50 Matches: 217
Percent Similarity: 67.37% Conservative: 103
Best Local Similarity: 45.68% Mismatches: 152
Query Match: 27.46% Indels: 3
DB: 11 Gaps: 2
US-09-667-170A-440 (1-2239) x Q9R0S5 (1-512)
Qy 310 GGAAGAGAGCGCTTTTCAGGAAGAGAGAGAAAGTCGAGCTGAAGAGGAAAGTCACTTTACTG 369
Db 24 GlyAspGlyAlaGlyProAlaAlaGluGlnValLysLeuLysLysGluIleSerLeuLeu 43
Qy 370 AGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGACGAGGAATCTTCTCATCTCTCTAAG 429
Db 44 AsnGlyValCysLeuIleValGlyAsnMetIleGlySerGlyIlePheValSerProLys 63
Qy 430 GGGCTGCTCCAGAACACGGCAGCGTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGG 489
Db 64 GlyValLeuMetTyrSerAlaSerPheGlyLeuSerLeuValIleTrpAlaValGlyGly 83
Qy 490 GTCCTGTCACTATTGGAGCTTTGTCTTATGCTGAATGGGAACAACATATAAGAAATCT 549
Db 84 IlePheSerValPheGlyAlaLeuCysTyrAlaGluLeuGlyThrIleLysLysSer 103
Qy 550 GGAGTTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTC 609
Db 104 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 123
Qy 610 TGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGA 669
Db 124 TrpThrSerLeuLeuIleIleGluProThrSerGlnAlaValIleAlaIleThrPheAla 143
Qy 670 CGCTACATCTGGAACCATTTTATTCAATGTGAAATCCCTGAACCTTGCATCAAGCTC 729
Db 144 AsnTyrMetValGlnProLeuPheProSerCysGlyAlaProTyrAlaAlaGlyArgLeu 163
Qy 730 ATTACAGCTGTGGGCATTAACCTGAGTGTGCTTAAATAGCATGAGTGTGAGCTGGAGC 789
Db 164 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly 183

QY 790 GCCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAATTATAGTC 849
Db 184 ThrLeuValGlnAspIlePheThrTyAlaLysValLeuAlaLeuIleAlaValIle 203
QY 850 CCTGGAGTTATGCAGCTAATTAAGTCAAAACGCAGAACTTTAAAGACGCCTTTTCAGGA 909
Db 204 AlaGlyIleValArgLeuGlyGlnGlyAlaThrThrAsnPheGluAspSerPheGluGly 223
QY 910 AGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATATGGAATGTATGATATGCT 969
Db 224 SerSerPheAlaMetGlyAspIleAlaLeuAlaLeuTySerAlaLeuPheSerTySer 243
QY 970 GGCTGGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAACCCCTGAAAAACCACTTCCC 1029
Db 244 GlyTrpAspThrLeuAsnTyValThrGluGluIleArgAsnProGluArgAsnLeuPro 263
QY 1030 CTTGCAATATGTATATCCATGGCCATTTGTACCAATTTGGCTATGTGTGACAAATGTGGCC 1089
Db 264 LeuSerIleGlyIleSerMetProIleValThrIleIleTyLeuLeuThrAsnValAla 283
QY 1090 TACTTTACGACCATTAATGCTGAGGAGCTGTGCTTCAAAATGCAGTGGCAGTGACCTTT 1149
Db 284 TyTrpSerValLeuAspIleLysAspIleLeuAlaSerAspAlaValAlaValThrPhe 303
QY 1150 TCTGACGGCTACTGGGAAATTTCTCATTAGCAGTCCGATCTTTTGTGCTCTCTCTGC 1209
Db 304 AlaAspGlnIlePheGlyIlePheAsnTrpThrIleProLeuAlaValAlaLeuSerCys 323
QY 1210 TTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGTCTCCAGGTTATTCTATGTTGCTCGA 1269
Db 324 PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuLeuPheValGlySerArg 343
QY 1270 GAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGAAGCACACTCTCTACCA 1329
Db 344 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 363
QY 1330 GCTGTATTGTTTGGACCCCTTTTGACAATGATAATGCTCTTCTGTGGAGACCTCGACAGT 1389
Db 364 SerLeuLeuPheAsnGlyIleLeuAlaLeuValTyLeuCysValGluAspIlePheGln 383
QY 1390 CTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTATTGGGTGGCAGTTGTGGGCTG 1449
Db 384 LeuIleAsnTyTrpSerPheSerTyTrpPhePheValGlyLeuSerIleValGlyGln 403
QY 1450 ATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCACTGTTCATC 1509
Db 404 LeuTyLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerLeuPhePhe 423
QY 1510 CCAGCTTTGTTTCTTCACATGCCTCTCATGTTGCTGTTGCCCTTCTCTATTCGGACCCA 1569
Db 424 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTySerAspThr 443
QY 1570 TTTAGTACAGGATGGCTTCGTATCATCTGACTGGAGTCCCTGCGTATTATCTCTTT 1629
Db 444 IleAsnSerLeuIleGlyIleGlyIleAlaLeuSerGlyLeuProPheTyPheLeuIle 463
QY 1630 ATT-----ATATGGGACAAGAACCCAGGTGGTTTAGAATAATGTACAGAGAAATAACC 1683
Db 464 IleArgValProGluHisLysArgProLeuCysLeuArgArgIleValAlaSerThrThr 483
QY 1684 AGAACATTACAAATAATA---CTGGAAGTTGTACCAGAGAAGAT 1725
Db 484 ArgTyLeuGlnIleIleCysMetSerValAlaAlaGluMetAsp 498
RESULT 11
Q9QZ66
ID Q9QZ66 PRELIMINARY; PRT; 512 AA.
AC Q9QZ66;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Amino acid transporter y+LAT1.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Ng A.M.L., Yao S.Y.M., Cheeseman C.I., Young J.D.;
RT "cDNA encoding rat jejunal amino acid transporter y+LAT1."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF200684; AAF07216.1; -
DR InterPro; IPR002293; AA/reel_prmease1.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
SQ SEQUENCE 512 AA; 55684 MW; 708905B2ACB2130B CRC64;
Alignment Scores:
Pred. No.: 6.2e-91 Length: 512
Score: 1131.50 Matches: 217
Percent Similarity: 67.37% Conservative: 103
Best Local Similarity: 45.68% Mismatches: 152
Query Match: 27.46% Indels: 3
DB: 11 Gaps: 2
US-09-667-170A-440 (1-2239) x Q9QZ66 (1-512)
QY 310 GGAAGAGAGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTTACTG 369
Db 24 GlyAspGlyAlaGlyProAlaAlaGluGlnValLysLeuLysGluIleSerLeuLeu 43
QY 370 AGGGAGTCTCCATTTATCATTGTCACCATCATTTGGCAGCAGGAATCTCTCATCTCTCTAAG 429
Db 44 AsnGlyValCysLeuIleValGlyAsnMetIleGlySerGlyIlePheValSerProLys 63
QY 430 GCGTGTCTCCAGAACACGCGGCGAGCGTGGCATGTCTCTGACCATCTGGACGGTGTGGG 489
Db 64 GlyValLeuMetTySerAlaSerPheGlyLeuSerLeuValIleTrpAlaValGlyGly 83
QY 490 GTCTGTCTACTATTGGAGCTTTTCTCTTATGCTGAATTTGGGAACAACATAAAGAAATCT 549
Db 84 IlePheSerValPheGlyAlaLeuCysTyAlaGluLeuGlyThrThrIleLysSer 103
QY 550 GGAGGTCAATTACACATATATTTTGAAGTCTTTTGGTCCATTACCAGCTTTTGTACGAGTC 609
Db 104 GlyAlaSerTyAlaTyIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 123
QY 610 TGGGTGGAACCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGA 669
Db 124 TrpThrSerLeuLeuIleIleGluProThrSerGlnAlaValIleAlaIleThrPheAla 143
QY 670 CGCTACATTTCTGGAACCACTTTTATTCAATGTGAATCCCTGAACCTTGGCATCAAGCTC 729
Db 144 AsnTyMetValGlnProLeuPheProSerCysGlyAlaProTyAlaAlaGlyArgLeu 163
QY 730 ATTACAGCTGTGGGCATAACTGTAGTGTGATGGTCTCTAAATAGCATGAGTGTGAGTGGAGC 789
Db 164 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyValLysTrpGly 183
QY 790 GCCCGGATCCAGATTTTCTTAACCTTTTTCAGAGCTCACAGCAATTTCTGATAATTATAGTC 849
Db 184 ThrLeuValGlnAspIlePheThrTyAlaLysValLeuAlaLeuIleAlaValIle 203
QY 850 CCTGGAGTTATGCAGCTAATTAAGGTCAAAACGCAGAACTTTAAAGACGCCTTTTCAGGA 909
Db 204 AlaGlyIleValArgLeuGlyGlnGlyAlaThrThrAsnPheGluAspSerPheGluGly 223
QY 910 AGACATTCAAGTATTACGCGGTTTGGCAGTGGCTTTTATTATGGAATGTATGATATGCT 969
Db 224 SerSerPheAlaMetGlyAspIleAlaLeuAlaLeuTySerAlaLeuPheSerTySer 243
QY 970 GGCTGGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAACCCCTGAAAAACCACTTCCC 1029
Db 244 GlyTrpAspThrLeuAsnTyValThrGluGluIleArgAsnProGluArgAsnLeuPro 263

QY	1030	CTTGCAATATGTATATCCATGCGCCATTTGTCACCATTTGGCTATGCTGACAAATGTGGCC	1089	RA	April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
Db	264	LeuSerIleGlyIleSerMetProIleValThrIleIleTyrLeuLeuThrAsnValAla	283	RA	Baliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
QY	1090	TACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCAGTACCTTT	1149	RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Db	284	TyTyrSerValLeuAspIleAspIleLeuAlaSerAspAlaValAlaValThrPhe	303	RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
QY	1150	TCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTCCGATCTTTGTTGCCCTCTCTGC	1209	RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Db	304	AlaAspGlnIlePheGlyIlePheAsnTrpThrIleProLeuAlaValAlaLeuSerCys	323	RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
QY	1210	TTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATTTCTATGTTGGCTCTCGA	1269	RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Db	324	PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuLeuPheValGlySerArg	343	RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
QY	1270	GAGGCTCACCTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCACACTCCCTACCA	1329	RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Db	344	GlUGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro	363	RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
QY	1330	GCTGTTATTGTTGCACCTTTTGACAAATGATAATGCTCTCTCTGGAGACCTCGACAGT	1389	RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Db	364	SerLeuLeuPheAsnGlyIleLeuAlaLeuValTyrLeuCysValGluAspIlePheGln	383	RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
QY	1390	CTTTTGAATTTCCAGGTTTGGCAGGCTTTTATTGGGCTGGCAGTTGCTGGGCTG	1449	RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Db	384	LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln	403	RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
QY	1450	ATTTATCTTCGATACAAATGCCAGATATGATGCTCTCTTCAAGTGCACACTGTTTCATC	1509	RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Db	404	LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerLeuPhePhe	423	RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
QY	1510	CCAGCTTTGTTTCCCTTCACATGCTCTTCATGTTGCTTCCCTTTCCTTATTCGGACCCA	1569	RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Db	424	ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr	443	RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
QY	1570	TTTAGTACAGGATGCTTGGCTTCATCATCTGCTGAGTGGCTGCTGCTTATCTCTTT	1629	RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Db	444	IleAsnSerLeuIleGlyIleGlyIleAlaLeuSerGlyLeuProPheTyrPheLeuIle	463	RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
QY	1630	ATT-----ATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTCAGAAATAACC	1683	RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Db	464	IleArgValProGluHisLysArgProLeuCysLeuArgArgIleValAlaSerThrThr	483	RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
QY	1684	AGAAATACAAATAATA---CTGGAAGTTGTACCAAGAAAGAT	1725	RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Db	484	ArgTyrLeuGlnIleIleCysMetSerValAlaAlaGluMetAsp	498	RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RESULT	12			RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Q9V9Y0				RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
ID	Q9V9Y0	PRELIMINARY;	PRT;	505 AA.	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
AC	Q9V9Y0;				Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
DT	01-MAY-2000 (TReMBLrel. 13, Created)				Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)				Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
DT	01-MAR-2002 (TReMBLrel. 20, Last annotation update)				"The genome sequence of Drosophila melanogaster.";
DE	CG1607 protein (Fragment).				Science 287:2185-2195(2000).
GN	CG1607.				EMBL; AE003778; AAF57148.1; -.
OS	Drosophila melanogaster (Fruit fly).				FlyBase; FBgn0039844; CG1607.
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				InterPro; IPR002293; AA/rel_prmease1.
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				InterPro; IPR002422; AA/rel_prmease2.
OC	Ephydroidea; Drosophilidae; Drosophila.				InterPro; IPR004841; Permease.
OX	NCBI_TaxID=7227;				Pfam; PF00324; aa_permeases; 1.
RN	[1]				NON_TER 505 505
RP	SEQUENCE FROM N.A.				SEQUENCE 505 AA; 55232 MW; CD8C0BD52529718C CRC64;
RC	STRAIN=BERKELEY;				
RX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				

Alignment Scores:
Pred. NO.: 2.38e-89 Length: 505
Score: 1113.50 Matches: 208
Percent Similarity: 64.50% Conservative: 110
Best Local Similarity: 42.19% Mismatches: 170
Query Match: 27.03% Indels: 5
DB: 5 Gaps: 2

US-09-667-170A-440 (1-2239) x Q9V9Y0 (1-505)

QY	235	AACGGGAGGCTGCCCTTCCCTGGGCAACAAGGAGCCACCTGGGCGAGGCGCTTTTCAGGA	294	QY	235	AACGGGAGGCTGCCCTTCCCTGGGCAACAAGGAGCCACCTGGGCGAGGCGCTTTTCAGGA
Db	11	SerSerArgLysSerSerIleValAsnGlyAsnGlyAspAlaSerAlaLysLeuThrAsn	30	Db	11	SerSerArgLysSerSerIleValAsnGlyAsnGlyAspAlaSerAlaLysLeuThrAsn
QY	295	AGAGACGCGCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAGAGAGAGAGAGAG	354	QY	295	AGAGACGCGCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAGAGAGAGAGAGAG
Db	31	GlyAspGlyAspGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	47	Db	31	GlyAspGlyAspGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly
QY	355	AAAGTCACCTTACTGAGGGGAGTCTCCATATTCATATTCATATTCATATTCATATTCAT	414	QY	355	AAAGTCACCTTACTGAGGGGAGTCTCCATATTCATATTCATATTCATATTCATATTCAT
Db	48	LysMetSerLeuLeuAsnGlyCysThrValIleValGlySerIleIleGlySerGlyIle	67	Db	48	LysMetSerLeuLeuAsnGlyCysThrValIleValGlySerIleIleGlySerGlyIle
QY	415	TTCATCTCTCTAAGGGCGTGTCTCCAGAACACAGGCGGCGGCGATGTCTCTGACCATC	474	QY	415	TTCATCTCTCTAAGGGCGTGTCTCCAGAACACAGGCGGCGGCGATGTCTCTGACCATC
Db	68	PheValSerProThrGlyValLeuMetTyrThrGlySerValAsnLeuAlaLeuVal	87	Db	68	PheValSerProThrGlyValLeuMetTyrThrGlySerValAsnLeuAlaLeuVal
QY	475	TGGACGGTGTGTGGGGTCTGTCTACATATTTGGAGCTTTTGTCTATGCTGAATTTGGGAACA	534	QY	475	TGGACGGTGTGTGGGGTCTGTCTACATATTTGGAGCTTTTGTCTATGCTGAATTTGGGAACA
Db	88	TrpValIleSerGlyLeuPheSerMetValGlyAlaTyrCysTyrAlaGluLeuGlyThr	107	Db	88	TrpValIleSerGlyLeuPheSerMetValGlyAlaTyrCysTyrAlaGluLeuGlyThr
QY	535	ACTATAAAGAAATCTGGAGGTCATTACACATATATTTTGGGAAGTCTTTGGTCCATTACCA	594	QY	535	ACTATAAAGAAATCTGGAGGTCATTACACATATATTTTGGGAAGTCTTTGGTCCATTACCA

Db 108 MetIleThrLysSerGlyAlaAspTyrAlaTyrIleMetGluThrPheGlyProPheMet 127
QY 595 GCTTTTACGAGTCTGGTGGAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATA 654
Db 128 AlaPheIleArgLeuTrpIleGluCysMetIleValArgProCysSerGlnAlaIleVal 147
QY 655 TCCCTGGCATTTGGAGCGCTACATTTCTGGAAACCAATTTTATTCAATGGAATCCCTGAA 714
Db 148 AlaLeuThrPheSerThrTyrValLeuLysProPhePheProGluCysThrProProGlu 167
QY 715 CTTGGCATCAAGCTCAATTACAGCTGTGGGCATTAAGTGTAGTATGGTCTTAATAGCATG 774
Db 168 AspSerAlaArgLeuLeuAlaValCysCysIleLeuValLeuThrLeuIleAsnCysTrp 187
QY 775 AGTGTACAGTGGAGCGCGCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAAT 834
Db 188 AspValLysTrpAlaThrAlaValGlnAspIlePheThrTyrAlaLysLeuLeuAlaLeu 207
QY 835 CTGATAAATTATAGTCCCTGGAGTTATGCAGCTAATTAAAGGTCAACGCGAGAACTTTAAA 894
Db 208 PheIleIleAlaThrGlyValTyrGlnLeuTyrLeuGlyAsnThrGlnTyrPhe--- 226
QY 895 GACGCTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGA 954
Db 227 ---ThrPheGluAsnThrAspThrLysValThrSerIleAlaLeuSerPheTyrSerGly 245
QY 955 ATGTATGCATATGCTGGTGGTGTACCTCAACTTTTGTACTGAAGAAGTAGAAACCCCT 1014
Db 246 LeuPheAlaTyrAsnGlyTrpAsnTyrLeuAsnPheIleIleGluGluLeuLysAspPro 265
QY 1015 GAAAAAACCATTTCCCTTGCATATGTATATATCCATGGCCATGTCCACCATTTGGTATGTG 1074
Db 266 ValLysAsnLeuProArgAlaIleAlaIleSerCysThrLeuValThrIleValTyrVal 285
QY 1075 CTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTCAAAATGCA 1134
Db 286 MetAlaAsnValSerPheTyrThrIleLeuSerProAspGluValMetGlySerSerAla 305
QY 1135 GTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTCCGATCTTT 1194
Db 306 ValAlaValThrTyrAlaGluArgAlaPheGlyMetLeuAlaTrpThrIleProValPhe 325
QY 1195 GTTGCCCTCTCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTC 1254
Db 326 ValAlaLeuSerThrPheGlyAlaValAsnGlyIleLeuLeuThrSerSerArgLeuPhe 345
QY 1255 TATGTTGGTCTCGAGAGGGTCACCTTCCAGAAATCTCTCCATGATCATGTCGCGAAG 1314
Db 346 TyrAlaGlyAlaAsnAsnGlyGlnMetProGluIleLeuThrMetIleGlnIleGlnArg 365
QY 1315 CACACTCTCTACCGCTGTTATTTGTTGACCCCTTTGACAAATGATAATGCTCTCTCTCT 1374
Db 366 PheThrProThrProAlaValLeuAlaMetAlaLeuLeuSerMetLeuTyrLeuThrVal 385
QY 1375 GGAGACCTCGACAGTCTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTG 1434
Db 386 SerAspIlePheAlaLeuIleAsnTyrValGlyPheAlaThrTrpLeuSerIleGlyVal 405
QY 1435 GCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTTCAAG 1494
Db 406 AlaValLeuCysLeuProTrpLeuArgTrpAlaGlnProAsnLeuProArgProIleArg 425
QY 1495 GTGCCACTGTTCAATCCAGCTTTGTTTCTTCATGCTTCATGCTTCATGGTGGCCTTTCC 1554
Db 426 ValProMetValPheProIleValTyrLeuIleAlaThrIlePheValThrValPro 445
QY 1555 CTCTATTCGGACCCATTTAGTACAGGATGGCTTCGTCATCACTCTGACTGGAGTCCCT 1614
Db 446 MetTyrAlaSerProValGluThrGlyTyrGlyIleLeuMetIleLeuSerSerIlePro 465
QY 1615 GCGTATTATCTCTTATTATATGGGCAAGAAACCCAGGTGGTTAGATAATATGTCAGAG 1674
Db 466 ValTyrLeuValPheIleAlaTrpLysAsnLysProIleThrPheGlnLysThrMetGly 485

QY 1675 AAAATAACACGACATTACAAATAATACTGGAAGTTGTA 1713
Db 486 GlyLeuThrGlnValLeuGlnLysLeuMetMetValVal 498
RESULT 13
Q9Y1A7
ID Q9Y1A7 PRELIMINARY; PRT; 499 AA.
AC Q9Y1A7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MND protein (LD25378p).
GN MND OR CG3297.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC Martin J.F., Hersperger E., Simcox A., Shearn A.;
RA "Minidiscs encodes a component of an endocrine signaling pathway
RT required for normal imaginal cell proliferation.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,

RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003532; AAF49688.1; -
 DR EMBL; AF139834; AAD39459.1; -
 DR EMBL; AY069533; AAL39678.1; -
 DR FlyBase; FBgn0002778; mnd.
 DR InterPro; IPR002293; AA/rel_prmeasel.
 DR InterPro; IPR004841; Permease.
 DR Pfam; PF00324; aa_permeases; 1.
 SQ SEQUENCE 499 AA; 54350 MW; E2882CC0737CC56F CRC64;

Alignment Scores:
 Pred. No.: 1.09e-88 Length: 499
 Score: 1106.00 Matches: 218
 Percent Similarity: 65.57% Conservative: 102
 Best Local Similarity: 44.67% Mismatches: 156
 Query Match: 26.84% Indels: 12
 DB: 5 Gaps: 3

US-09-667-170A-440 (1-2239) x Q9YIA7 (1-499)

QY 265 GAGCCACCTGGCAGGACGCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGACGCGCTTT 324
 Db 22 GluProAsnAsnSerThrAlaAspSerGlySerGlnGly----- 34

QY 325 TCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATT 384
 Db 35 SerGly-----ValLysLeuLysLysGlnIleGlyLeuLeuAspGlyValAlaIle 51

QY 385 ATCATTTGGCACCACATATGGAGCAGGAATCTTCATCTCTCCCTAAGGCGTGTCCAGAAC 444
 Db 52 IleValGlyValIleValGlySerGlyIlePheValSerProLysGlyValLeuLysPhe 71

QY 445 ACGGCGAGCGTGGCAATGCTCTGACCATCTGGACGGTGTGTGGGCTCTGTCACATATT 504
 Db 72 SerGlySerIleGlyGlnSerLeuIleValTrpValLeuSerGlyValLeuSerMetVal 91

QY 505 GGAGCTTGCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGAGGTCATTACACA 564
 Db 92 GlyAlaLeuCysTyrAlaGluLeuGlyThrMetIleProLysSerGlyGlyAspTyrAla 111

QY 565 TATATTTTGGAAAGTCTTTGGFCCATTACAGCTTTTGTACGAGTCTGGTGGAACTCCTC 624
 Db 112 TyrIleGlyThrAlaPheGlyProLeuProAlaPheLeuTyrLeuTrpValAlaLeuLeu 131

QY 625 ATAATACGCGCTGACGCTACTGCTGTGATATCCCTGGCATTCTGGACGCTACATCTGGA 684
 Db 132 IleLeuValProThrGlyAsnAlaIleThrAlaLeuThrPheAlaIleTyrLeuLys 151

QY 685 CCATTTTATCAATGTGAATCCCTGAACCTTGCATCAAGCTCATTACAGCTGTGGC 744
 Db 152 PropheTrpProSerCysAspAlaProIleGluAlaValGlnLeuLeuAlaAlaMet 171

QY 745 ATAAGCTAGTGTCTTAAATAGCATGAGTGTGAGTGTGAGCGCGCCGCGATCCAGATT 804
 Db 172 IleCysValLeuThrLeuIleAsnCysTyrAsnValLysTrpValThrArgValThrAsp 191

QY 805 TTCTTAACCTTTGCAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGCAG 864
 Db 192 IlePheThrGlyThrLysValValAlaLeuLeuValIleValGlyAlaGlyValTrpTrp 211

QY 865 CTAATTAAGGTCACAGCAACCTTAAAGACGCGCTTTTCAGGAAGAGATTCAAGTATT 924
 Db 212 LeuPheAspGlyAsnThrGluHisTrpAspAsnProPheSerGlyGlyLeuGlnAspPro 231

QY 925 ACGCGGTTGCCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTC 984
 Db 232 GlyTyrIleAlaLeuAlaPheTyrSerGlyLeuPheSerTyrSerGlyTrpAsnTyrLeu 251

QY 985 AACTTTGTTACTGAAGAGTAGAAACCCCTGAAAAACCATTCCTTCCCTTGCATATATATA 1044
 Db 252 AsnPheValThrGluGluLeuLysAspProTyrArgAsnLeuProLysAlaIleCysIle 271

QY 1045 TCCATGGCCATGTCCACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCAT 1104
 Db 272 SerMetProValValThrValIleTyrMetIleThrAsnIleAlaTyrPheSerValLeu 291

QY 1105 AATGCTGAGGAGCTGCTGCTTTCAATGTCAGTGGCAGTGCCTTTCTGAGCGGCTACTG 1164
 Db 292 SerProAspGluIleLeuSerSerAspAlaValAlaValThrPheGlyAspLysMetLeu 311

QY 1165 GGAATTTCTCATTAGCAGTTCGATCTTTGTTGTCCTCTCTGCTTTGGCTCCCATGAAC 1224
 Db 312 GlyTyrMetSerTrpIleMetProPheAlaValAlaCysSerThrPheGlySerLeuAsn 331

QY 1225 GGTGGTGTGTTGCTCTCCAGGTTATCTATGTTGCGTCTCGAGAGGGTCACCTTCCA 1284
 Db 332 GlyAlaIlePheAlaSerSerArgLeuPhePheValGlyAlaArgAsnGlyHisLeuPro 351

QY 1285 GAAATCCTCTCCATGATTCATGTCGCAAGCAGCAGTCTCTACAGCTGTTATTGTTTG 1344
 Db 352 AlaAlaIleSerLeuIleAsnValAsnCysLeuThrProValProSerLeuIlePheLeu 371

QY 1345 CACCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTGAATTTCTC 1404
 Db 372 GlyValLeuThrLeuLeuLeuPheIleGluAspValTyrValLeuIleAsnTyrVal 391

QY 1405 AGTTTTCAGGCTGGCTTTTATTGGGCTGGCAGTGTGCTGGGCTGATTATCTTCGATAC 1464
 Db 392 SerTyrValGluAlaLeuPheThrLeuIleSerValSerGlyLeuLeuTrpMetArgTyr 411

QY 1465 AAATGCCAGATATGCATCGCTCTTCAAGGTGGCAGTGTTCATCCAGCTTTGTTTCC 1524
 Db 412 LysGlnProLysThrGluArgProIleLysValAsnLeuAlaLeuProIleIleTyrLeu 431

QY 1525 TTCACATGCTCTTTCATGCTGGCTGCTCTTCCCTCTATTCGGACCCATTTAGTACAGG 1584
 Db 432 IleValCysLeuPheLeuValIlePheSerCysThrGlnThrProTyrValValGlyIle 451

QY 1585 GGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATATCTCTTTTATTATATGGGACAAG 1644
 Db 452 GlyThrIleIleLeuSerGlyIleProValTyrTyrLeuThrIle-----HisLys 469

QY 1645 AAACCCAGGTGGTTTAGAATAATGTCAGAGAAAATAACAGAACATTAACAATAACTG 1704
 Db 470 ProValLysTrpLeuAlaAspThrSerGlnAlaIleAsnLeuTrpCysSerLysPhePhe 489

QY 1705 GAAGTTGTACCAAGAGAGATAAG 1728
 Db 490 IleCysMetProAsnGlnGluLys 497

RESULT 14
 Q9N1Q4 PRELIMINARY; PRT; 535 AA.
 ID Q9N1Q4
 AC Q9N1Q4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 4F21c-5.
 GN 4F21c-5.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20473842; PubMed=10631289;
 RA Rajan D.P., Kekuda R., Huang W., Devoe L.D., Leibach F.H.,
 RA Prasad P.D., Ganapathy V.;
 RT "Cloning and functional characterization of a Na(+)-independent,
 RT broad-specific neutral amino acid transporter from mammalian
 RT intestine";
 RL Biochim. Biophys. Acta 1463:6-14 (2000).
 DR EMBL; AF170106; AAF26282.1; -
 DR InterPro; IPR002293; AA/rel_prmeasel.

DR InterPro: IPR004760; L_AA_transport.
DR InterPro: IPR004841; Permease.
DR Pfam: PF00324; aa_permeases; 1.
DR TIGRFAMS: TIGR00911; 2A0308; 1.
SQ SEQUENCE 535 AA; 58321 MW; D4939C0F971CB190 CRC64;

Alignment Scores:

Pred. No.: 1.03e-87 Length: 535
Score: 1095.00 Matches: 216
Percent Similarity: 65.97% Conservative: 98
Best Local Similarity: 45.38% Mismatches: 154
Query Match: 26.58% Indels: 8
DB: 6 Gaps: 3

US-09-667-170A-440 (1-2239) x Q9N1Q4 (1-535)

QY 316 GACGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGGA 375
Db 26 GluAlaSerSerGlyGlyGlyValAlaLeuLysLysGluIleGlyLeuValSerAla 45
QY 376 GTCTCCATTATATGTCACCATCATTTGGAGCAGGAATCTTCACTCTCTCCTAAGGGCGTG 435
Db 46 CysGlyIleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyVal 65
QY 436 CTCAGAACACGGCAGCGTGGGCATGTCTGTGACCATCTGGACGCTGTGGGTCTCTG 495
Db 66 LeuGluAsnAlaGlySerValGlyLeuAlaValIleValTrpIleValThrGlyLeuIle 85
QY 496 TCACTATTGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGT 555
Db 86 ThrAlaValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGly 105
QY 556 CATTACACATATATTTTGAAGTCTTTGGTCCATTACAGCTTTTGTACGAGCTCTGGGTG 615
Db 106 AspTyrSerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIle 125
QY 616 GAACCTCCTCAATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGCTAC 675
Db 126 AlaValLeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyr 145
QY 676 ATTCTGGAAACCATTTTATTCAATGTGAATCCCTGAACTGGCATCAAGCTCATACATA 735
Db 146 ValLeuGlnProLeuPheProThrCysPheProProAspSerGlyLeuArgLeuLeuAla 165
QY 736 GCTGTGGGATAAAGTGTAGTGTGCTCTAAATAGCATAGTGTGAGTGGAGCGCCGG 795
Db 166 AlaIleCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 185
QY 796 ATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGA 855
Db 186 ValGlnAspIlePheThrAlaGlyLysLeuLeuAlaLeuAlaLeuIleIleMetGly 205
QY 856 GTTATGAGCTAATTAAGGT-----CAAACGCAGAACTTTAAAGACGCC 900
Db 206 ValValGlnIleCysLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 925
QY 901 TTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTAT 960
Db 226 Phe-----GlnGluProAspIleGlyLeuIleAlaLeuAlaPheLeuGlnGlySerPhe 243
QY 961 GCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCTGAAAAA 1020
Db 244 AlaTyrGlyGlyTyrAsnPheLeuAsnTyrValThrGluLeuValAspProTyrLys 263
QY 1021 ACCATTCCCTTGTCAATATATATATCCATGCGCCATTGTACCATTTGGCTATGTGCTGACA 1080
Db 264 AsnLeuProArgAlaIlePheIleSerIleProLeuValThrPheValThrPheAla 283
QY 1081 AATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAATGTCAGTGGCA 1140
Db 284 AsnValAlaTyrIleThrAlaMetSerProGlnGluLeuAlaSerAsnAlaValAla 303
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTGTGGC 1200

Db 304 ValThrPheGlyGluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAla 323
QY 1201 CTCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAGGTTATTTCTATGTT 1260
Db 324 LeuSerThrPheGlyGlyValAsnGlySerLeuPheThrSerArgLeuPhePheAla 343
QY 1261 GCGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGCAAGCAGACT 1320
Db 344 GlyAlaArgGluGlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThr 363
QY 1321 CCTCTACAGCTGTTATTGTTTTCACCCCTTTTGACAATGATGCTCTCTCTGAGAGAC 1380
Db 364 ProIleProAlaLeuLeuPheThrCysLeuSerThrLeuLeuMetLeuValThrSerAsp 383
QY 1381 CTCGACAGCTCTTTTGAATTTCTCAGTTTTCGCCAGGTGGCTTTTATTGGCTGGCAGTT 1440
Db 384 MetTyrThrLeuIleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrVal 403
QY 1441 GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGTCATCGTCTTCAAGGTGCCA 1500
Db 404 AlaGlyGlnIleValLeuArgTrpLysLysProAspIleProArgProIleLysIleAsn 423
QY 1501 CTGTTTCATCCAGCTTTTGTCTTCTTCCATGCTCTTCCATGTTGCTTTCCTCTAT 1560
Db 424 LeuLeuPheProIleIleTyrLeuLeuPheTrpAlaPheLeuLeuIlePheSerLeuTrp 443
QY 1561 TCGGACCCATTTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTAT 1620
Db 444 SerGluProValValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyr 463
QY 1621 TATCTCTTTTATATATGCGACAAAGAACCCAGGTTGTTAGATAATATGTCAGAGAAAAATA 1680
Db 464 PheLeuGlyValTyrTrpGlnHisLysProLysCysPheAsnAspPheIleGluLeuLeu 483
QY 1681 ACCAGAACATTACAAATAATACTGGAAGTTGTA--CCAGAAGAAGAT 1725
Db 484 ThrLeuValSerGlnLysMetCysValValValTyrProGluValAsp 499

RESULT 15
Q9VB75
ID Q9VB75 PRELIMINARY; PRT; 517 AA.
AC Q9VB75;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG5070 protein (GH08870P).
GN CG5070.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt R.G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balleson R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]

SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003759; AAF56667.1; -;
DR EMBL; AY060663; AAL28211.1; -;
DR FlyBase; FBgn0039487; CG6070.
DR InterPro; IPR002293; AA/rel_prmeasel.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
SQ SEQUENCE 517 AA; 56522 MW; F1BDB0A7F51F18C8 CRC64;

Alignment Scores:
Pred. No.: 2,3e-87 Length: 517
Score: 1091.00 Matches: 210
Percent Similarity: 64.86% Conservative: 113
Best Local Similarity: 42.17% Mismatches: 167
Query Match: 26.48% Indels: 8
DB: 5 Gaps: 4

US-09-667-170A-440 (1-2239) x Q9VB75 (1-517)

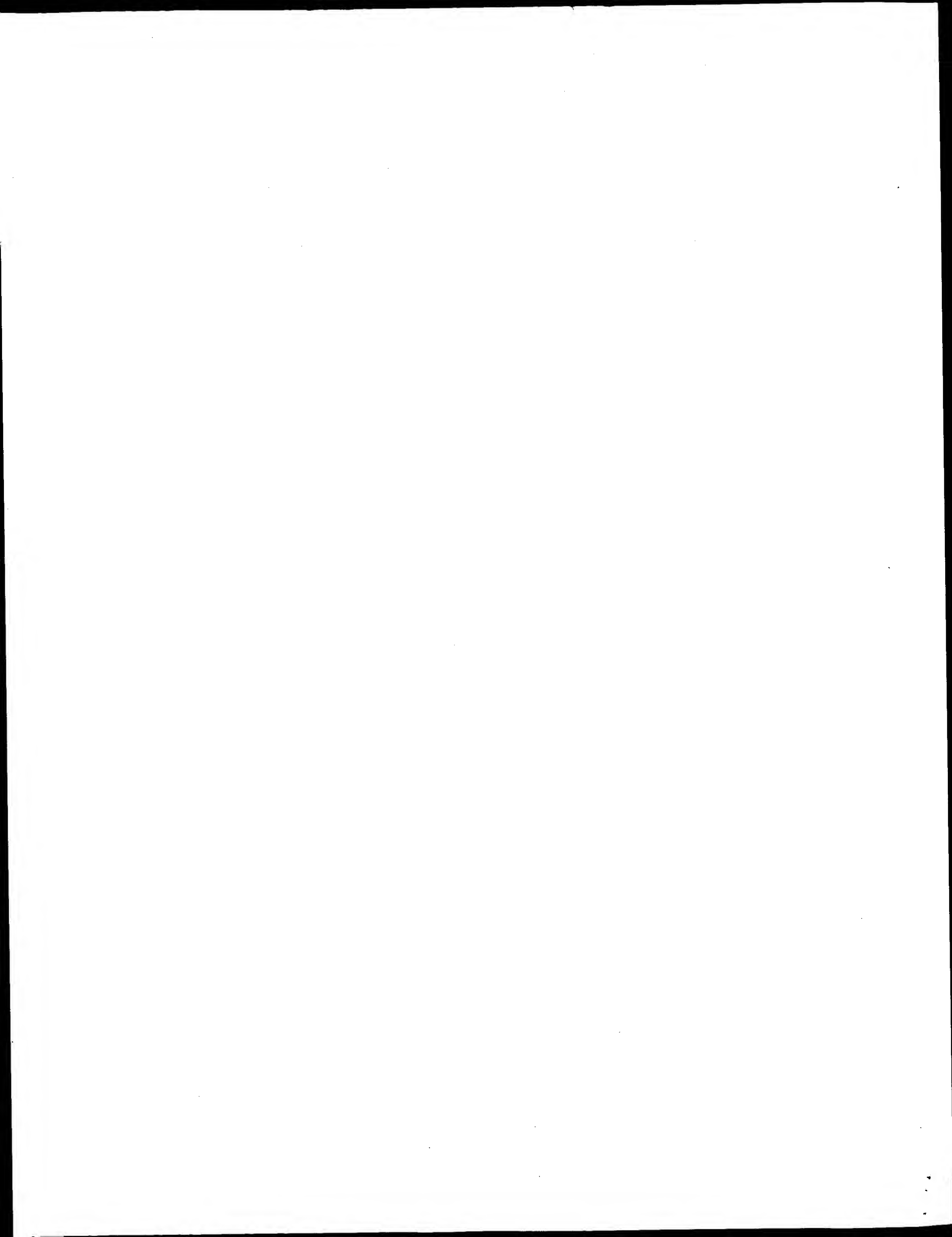
QY 247 CCTTCCCTGGGCAACAGAGCCACCT-----GGGCAGGACGCCCTTTTCAGGAAGA 297
Db 19 ProSerSerGlyAsn-----ProProSerSerIleAlaGlnProAlaAlaAsnProAla 36
QY 298 GACGCCCTTTTCAGGAAGAGAC-----GCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAG 351
Db 37 GluLysAlaGlnCysArgGluGlySerAlaGluSerSerArgValValIleLys 56
QY 352 AGGAAAGTCACTTTACTAGGGGAGTCTCCATTATCATTTGGCACCATTATGGAGCAGGA 411
Db 57 LysGlnLeuGlyLeuLeuGluGlyValAlaIleLeuGlyIleIlePheGlySerGly 76
QY 412 ATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGCAGCGTGGCGATGCTCTGACC 471
Db 77 IlePheValSerProLysGlyValIleArgGluValGluSerValGlyAlaSerLeuVal 96
QY 472 ATCTGGACGGTGTGTGGGGTCTCTGCTCACTATTTGGAGCTTTGTCTTATGCTGAATGGGA 531
Db 97 IleTrpValLeuCysGlyLeuLeuSerMetIleGlyAlaLeuCysTyrAlaGluLeuGly 116
QY 532 ACACTATAAGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGGTCCATTA 591
Db 117 ThrAlaIleProLysSerGlyGlyAspTyrAlaTyrIlePheGluAlaTyrGlySerLeu 136

QY 592 CCAGCTTTGTACGAGTCTGGTGGGAACCTCCTCATAATACGCCCTGCAGCTACTGCTGTG 651
Db 137 ProAlaPheLeuTyrLeuTrpAspAlaMetMetIlePheValProThrThrAsnAlaIle 156
QY 652 ATATCCCTGGCATTTGGACGCTACATTTCTGGAACCACTTTT---ATTCAATGTGAATC 708
Db 157 MetGlyLeuThrPheAlaSerTyrValLeuGluProPhePheGlyGlyAlaCysGluIle 176
QY 709 CCTGAACCTTGCATCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGATGCTCTAAT 768
Db 177 ProLysIleAlaLeuGlnLeuLeuAlaAlaIleThrIleCysPheLeuThrTyrLeuAsn 196
QY 769 AGCATGAGTGTGAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGCAGGCTCACA 828
Db 197 SerTyrTyrMetLysValThrThrLysMetGlnAsnValIleMetPheThrLysIleAla 216
QY 829 GCAATTCGTATATATAGTCCCTGGAGTATGCTAGTATGCTAGTAAAGTCAAACGAGAAC 888
Db 217 AlaLeuValLeuIleIleLeuValGlyLeuValTrpMetMetMetGlyAsnValGluAsn 236
QY 889 TTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTAT 948
Db 237 PheThrArgProPheAspAsnThrGluThrAspProGlyLysMetSerValAlaPheTyr 256
QY 949 TATGGAATGTATGCATATGCTGGTGGTGTCTTACCTCAACTTTTGTACTGAAGTAGAA 1008
Db 257 SerGlyIlePheSerTyrAlaGlyTrpAsnTyrLeuAsnPheMetThrGluGluLeuArg 276
QY 1009 AACCTTGAAAAACCACTTCCCTTGCATATGATATATATCCATGGCCATTTGTCACCAT 1068
Db 277 AspProTyrArgAsnLeuProArgAlaIleTyrIleSerLeuProLeuValThrGlyIle 296
QY 1069 TATGTGCTGACAAATGTGGCCTACTTTACGACCACTTAATGCTGAGGAGCTGCTTCTCA 1128
Db 297 TyrValLeuAlaAsnValAlaTyrLeuAlaValLeuSerProSerGluMetIleAlaSer 316
QY 1129 AATGCAGTGGCAGTACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCG 1188
Db 317 AsnAlaIleAlaValThrPheGlyAspLysIleLeuGlyValPheSerLeuIleIlePro 336
QY 1189 ATCTTTGTGGCCTCTCCTGCTTTGGCTCCATGACGGTGGTGTGTTGCTGCTCTCCAGG 1248
Db 337 LeuMetValAlaIleSerAlaPheGlyGlyLeuSerValHisIleMetThrSerSerArg 356
QY 1249 TTATTTCTATGTTGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTC 1308
Db 357 IleCysPheValGlyAlaArgAsnGlyHisMetProAlaIleLeuSerHisIleSerVal 376
QY 1309 CGCAAGCACACTCCTCTACCACTGTTTATTTGTCACCCCTTTTGACAATGATAATGCTC 1368
Db 377 LysSerTyrThrProLeuProSerLeuValPheLeuCysPheLeuSerIleValMetLeu 396
QY 1369 TTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCTCAGTTTGGCAGGTGGCTTTTATT 1428
Db 397 ValValSerAspValTyrValLeuIleThrTyrAlaSerIleValGluSerPhePheIle 416
QY 1429 GGGCTGGCAGTGTGGCTGCTGCTATCTTCGATACAAATGCCAGATATGTCATGCTCCT 1488
Db 417 MetLeuSerValSerAlaValLeuTyrPheArgTyrThrArgProCysMetGluArgPro 436
QY 1489 TTCAGGTGCTCACTGTTTCATCCAGCTTTGTTTCTTCCATCAGATGCCTCTTCACTGTTGCC 1548
Db 437 IleLysValAlaMetTrpIleProAlaLeuPheValIleValCysAlaPheLeuValVal 456
QY 1549 CTTTCCCTCTATTCGGACCCCAATTTAGTACAGGGATTGGCTGCTCATCTACTGCTGGA 1608
Db 457 ValProIleTyrValAlaProTyrGluValGlyMetGlyValLeuIleThrIleIleGly 476
QY 1609 GTCCCTGCTATTATCTCTTTATATATATGGACAAGAACCCAGGTGGTTTAGAATAATG 1668
Db 477 IleProPheTyrTyrValGlyValValTrpLysAsnLysProLysTrpValGlnSerThr 496
QY 1669 TCAGAGAAAAATAACCAACAATTTACAAATAATACTGGAGTGTGTACCAGAGAA 1722

us-09-667-170a-440.n2p.rspt

Db 497 IleAspSerValThrPheThrCysGlnLysLeuPheMetSerAlaLysGluGlu 514

Search completed: April 16, 2003, 16:14:05
Job time : 113.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 10:55:43 ; Search time 93 Seconds
(without alignments)
7383.325 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	1.2	737	2	US-08-861-549-2
2	24	1.1	40000	4	US-09-780-049-18
3	24	1.1	162450	4	US-09-345-882-1
4	23	1.0	72928	3	US-09-009-913-1
5	22	1.0	333	4	US-09-018-584A-27
6	22	1.0	350	2	US-08-623-906A-17
7	22	1.0	11288	4	US-08-646-301A-1
8	22	1.0	11288	4	US-08-481-968A-4
9	22	1.0	11288	4	US-08-154-712B-4
10	22	1.0	36651	4	US-09-738-894A-3
11	22	1.0	246240	2	US-08-724-394A-20
12	22	1.0	246240	2	US-08-724-394A-21
13	22	1.0	246240	2	US-08-724-394A-22
14	21	0.9	544	4	US-09-280-116-247
15	21	0.9	579	4	US-09-328-111-68
16	21	0.9	668	4	US-09-328-111-687
17	21	0.9	704	4	US-09-475-316A-26
18	21	0.9	950	3	US-08-984-550-1
19	21	0.9	2636	4	US-09-370-807-5
20	21	0.9	2636	4	US-09-921-259-5
21	21	0.9	5510	1	US-08-123-161A-7
22	21	0.9	5510	1	US-08-483-278-7
23	21	0.9	5852	1	US-07-867-106-2
24	21	0.9	8353	3	US-08-611-587-1
25	21	0.9	22067	4	US-09-820-001-3
26	21	0.9	36651	4	US-09-738-894A-3
27	21	0.9	84495	4	US-09-797-906-3

28	20	0.9	985	4	US-08-842-306B-1	Sequence 1, Appli
29	20	0.9	985	4	US-08-838-973B-1	Sequence 1, Appli
30	20	0.9	985	4	US-08-771-212A-1	Sequence 1, Appli
31	20	0.9	1229	1	US-08-440-861-1	Sequence 1, Appli
32	20	0.9	1229	1	US-08-433-854-1	Sequence 1, Appli
33	20	0.9	1229	1	US-08-174-745A-1	Sequence 1, Appli
34	20	0.9	1229	2	US-08-433-885-1	Sequence 1, Appli
35	20	0.9	1229	2	US-08-433-908B-1	Sequence 1, Appli
36	20	0.9	1229	4	US-08-410-614-1	Sequence 1, Appli
37	20	0.9	1242	4	US-08-413-974-1	Sequence 1, Appli
38	20	0.9	1242	4	US-08-434-418-1	Sequence 1, Appli
39	20	0.9	1242	4	US-08-433-288-1	Sequence 1, Appli
40	20	0.9	1242	4	US-08-174-739A-1	Sequence 1, Appli
41	20	0.9	1242	4	US-08-434-256-1	Sequence 1, Appli
42	20	0.9	1304	4	US-09-372-422A-15	Sequence 15, Appl
43	20	0.9	1400	4	US-09-018-584A-35	Sequence 35, Appl
44	20	0.9	1460	4	US-09-257-179-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-861-549-2
; Sequence 2, Application US/08861549
; Patent No. 5874246
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Zhang, Hong
; TITLE OF INVENTION: NEW TRANSCRIPTION REPRESSOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,549
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0306 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAB01
; CLONE: 1620089
US-08-861-549-2

Query Match 1.2%; Score 27; DB 2; Length 737;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97099..97145
; OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97099..97145
; OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97130..97177
; OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97130..97177
; OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99075..99121
; OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99075..99121
; OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99094..99140
; OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99094..99140
; OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103783..103828
; OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103783..103828
; OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106918..106966
; OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106918..106966
; OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58

```

```

; FEATURE:
; NAME/KEY: allele
; LOCATION: 108084..108130
; OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108084..108130
; OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
; FEATURE:

```

```

Query Match 1.1%; Score 24; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 27 CCAGCCTGGGTGACAGTGAGACTC 50
|||||
Db 54003 CCAGCCTGGGTGACAGTGAGACTC 53980

```

RESULT 4

```

US-09-009-913-1/c
; Sequence 1, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: Axys Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3231
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72928 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-009-913-1

```

```

Query Match 1.0%; Score 23; DB 3; Length 72928;
Best Local Similarity 100.0%; Pred. No. 1.1;

```

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAGTTGAAGTACGACAGATCA 24
 |||||
 Db 47805 GAGTTGAAGTACGACAGATCA 47783

RESULT 5

US-09-018-584A-27
 ; Sequence 27, Application US/09018584A
 ; Patent No. 6238863
 ; GENERAL INFORMATION:
 ; APPLICANT: Schumm, James W.
 ; APPLICANT: Bacher, Jeffery W.
 ; TITLE OF INVENTION: MATERIALS AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFYING AND ANALYZING
 ; TITLE OF INVENTION: REPEAT DNA MARKERS
 ; NUMBER OF SEQUENCES: 147
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Promega Corporation
 ; STREET: 2800 Woods Hollow Road
 ; CITY: Madison
 ; STATE: Wisconsin
 ; COUNTRY: U.S.A.
 ; ZIP: 53711-5399
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
 ; COMPUTER: IBM compatible PC
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: Word 97 (DOS text format)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/018,584A
 ; FILING DATE: 04-Feb-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Grady J. Frenchick
 ; REGISTRATION NUMBER: 29,018
 ; REFERENCE/DOCKET NUMBER: 16026.9180
 ; TELEPHONE: (608) 257-3501
 ; TELEFAX: (608) 257-2275
 ; INFORMATION FOR SEQ ID NO: 27:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 333 bp
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Circular
 ; MOLECULE TYPE: Genomic DNA
 ; HYPOTHETICAL: no
 ; IMMEDIATE SOURCE:
 ; LIBRARY: plasmid, pGem3zf(+)
 ; CLONE: G539
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT: 15q26.2
 ; US-09-018-584A-27

Query Match 1.0%; Score 22; DB 4; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4.5;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AAAAAAGAAAGAAAGAAAGA 92
 |||||
 Db 58 AAAAAAGAAAGAAAGAAAGA 79

RESULT 6

US-08-623-906A-17
 ; Sequence 17, Application US/08623906A
 ; Patent No. 5874217
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevenson, Tamara
 ; APPLICANT: Dvorak, Jan
 ; APPLICANT: Halverson, Joy

; TITLE OF INVENTION: Microsatellite Sequences for Canine
 ; TITLE OF INVENTION: Genotyping
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/623,906A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sherwood, Pamela J.
 ; REGISTRATION NUMBER: 36,677
 ; REFERENCE/DOCKET NUMBER: A-62282/BIR
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-781-1989
 ; TELEFAX: 415-398-3249
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 350 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..109
 ; OTHER INFORMATION: /note= "Nucleotides 1-109 are
 ; OTHER INFORMATION: unique flanking sequence"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 110..195
 ; OTHER INFORMATION: /note= "Nucleotides 110-195 are
 ; OTHER INFORMATION: repeat sequence"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 196..350
 ; OTHER INFORMATION: /note= "Nucleotides 196-350 are
 ; OTHER INFORMATION: unique flanking sequence"
 ; US-08-623-906A-17

Query Match 1.0%; Score 22; DB 2; Length 350;
 Best Local Similarity 100.0%; Pred. No. 4.5;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AAAAAAGAAAGAAAGAAAGA 92
 |||||
 Db 121 AAAAAAGAAAGAAAGAAAGA 142

RESULT 7

US-08-646-301A-1
 ; Sequence 1, Application US/08646301A
 ; Patent No. 6194211
 ; GENERAL INFORMATION:
 ; APPLICANT: Richards, Cynthia Ann
 ; APPLICANT: Huber, Brian E.
 ; TITLE OF INVENTION: Transcriptional Regulatory Sequence of Carcinoembryonic
 ; Patent No. 6194211
 ; TITLE OF INVENTION: Antigen for Expression Targeting
 ; FILE REFERENCE: PB1508USW
 ; CURRENT APPLICATION NUMBER: US/08/646,301A
 ; CURRENT FILING DATE: 1996-05-16
 ; NUMBER OF SEQ ID NOS: 25

	Best Local Similarity	100.0%;	Pred. No.	2.6;	
	Matches	22; Conservative	0; Mismatches	0; Indels	0; Gaps
QY	28	CAGCCTGGGTGACAGTGAGACT	49	.	
ph	236746	CAGCCTGGGTGACAGTGAGACT	236767		

RESULT 13
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237

RESULT 13
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereeto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
;

NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/724,394A
 FILING DATE: 01-OCT-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Fitts, Renee A.
 REGISTRATION NUMBER: 35,136
 REFERENCE/DOCKET NUMBER: 017957-000100
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 415-576-0300
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 246240 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1..246240
 OTHER INFORMATION: /note= "HLA-H.CONTIG"
 US-08-724-394A-22

Query Match	1.0%;	Score 22;	DB 2;	Length 246240;
Best Local Similarity	100.0%;	Pred. No. 2.6;		
Matches 22;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 28	CAGCCTGGGTGACAGTGAGACT	49		
Db 236746	CAGCCTGGGTGACAGTGAGACT	236767		

RESULT 14
US-09-280-116-247
; Sequence 247, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:

Query Match
1.0%; Score 22; DB 2; Length 246240;

Search completed: April 16, 2003, 14:34:09
Job time : 1441 secs

APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035800/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
CURRENT FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 247
LENGTH: 544
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: hemoglobinase
NAME/KEY: misc_feature
LOCATION: (1)..(544)
OTHER INFORMATION: n = a, t, c or g
US-09-280-116-247

Query Match 0.9%; Score 21; DB 4; Length 544;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 AGTGAGACTCTGTCTCAAAACA 61
|||||
Db 370 AGTGAGACTCTGTCTCAAAACA 390

RESULT 15
US-09-328-111-68
Sequence 68, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 68
LENGTH: 579
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(579)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-68

Query Match 0.9%; Score 21; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2219 TTTATTAAAAA 2239
|||||
Db 188 TTTATTAAAAA 208

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      April 16, 2003, 13:16:30 ; Search time 231 Seconds
            (without alignments)      9736.858 Million cell updates/sec
```

```

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggagggttgagtgagcagaq.....ttattaaaaa
2230

```

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: . 639749 seqs, 502280978 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

```
Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

SUMMARIES

SUMMARIES						
Result No.	Score	Query %		Length	DB	ID
		Match	Description			
1	2239	100.0	2239	9	US-09-854-133-440	Sequence 440, App
2	2239	100.0	2239	10	US-09-738-973-440	Sequence 440, App
3	1867	83.4	5981	9	US-09-854-133-441	Sequence 441, App
4	1867	83.4	5981	10	US-09-738-973-441	Sequence 441, App
5	1816	81.1	3144	9	US-10-163-866-31	Sequence 31, Appl
6	1722	76.9	2482	9	US-10-163-866-29	Sequence 29, Appl
7	1454	64.9	1861	9	US-10-163-866-30	Sequence 30, Appl
8	1420	63.4	1542	9	US-10-163-866-33	Sequence 33, Appl
9	1339	59.8	2000	9	US-10-163-866-34	Sequence 34, Appl
10	1299	58.0	1528	9	US-10-163-866-52	Sequence 52, Appl
11	868	38.8	1268	9	US-10-163-866-53	Sequence 53, Appl
12	415	18.5	520	9	US-10-163-866-32	Sequence 32, Appl
13	281	12.6	337	9	US-09-854-133-442	Sequence 442, App
14	281	12.6	337	10	US-09-738-973-442	Sequence 442, App
15	107	4.8	572	9	US-10-046-935-1307	Sequence 1307, Ap
16	107	4.8	572	9	US-09-878-178-1307	Sequence 1307, Ap
17	107	4.8	572	9	US-10-146-502-1307	Sequence 1307, Ap
18	32	1.4	261	10	US-09-998-598-2201	Sequence 2201, Ap
19	32	1.4	2440	10	US-09-962-436-286	Sequence 286, App

ALIGNMENTS

RESULT 1

US-09-854-133-440
; Sequence 440, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:

; APPLICANT: Lodes, Michael J.

; APPLICANT: Mohamath, Raodoh

APPLICANT: Henderson, Robert A.

; APPLICANT: Benson, Darin R.

; APPLICANT: Secrist, Heather

;	TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
:	TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121 475010

FILE REFERENCE: 210121.475C10
CURRENT APPLICATION NUMBER: 11

: CURRENT FILING DATE: 2001-05-11
 : CURRENT APPLICATION NUMBER: US/09/854,133

;; CURRENT FILING DATE: 2001-05-11
;; NUMBER OF SEQ ID NOS: 735

```

; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Win

```

```

/ SOFTWARE: FASTCSEQ FOR WINDOWS VERSION 3.0
; SEQ ID NO 440

```

; LENGTH: 2239

TYPE: DNA

ORGANISM:

US-09-854-133-440

Query Match	100.0%;	Score 2239;	DB 9;	Length 2239;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2239; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

1 GGAGGTTGAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAC 60

[illegible]

61 AGAATTAAAGAAAAAGAAAGAAAAAGAGAGAGGAAATTCACGGCCAAATTGTGG 120

121 CATAGATTTATCATATCTGGATTTTGGATTCCTTTGTTTTCTCATCTACTGGATCA 180

QY	1801	TTTCTGAAAGCTAGAGAATTACAACCTTTGGTGATAAACCAAAAGGAGTCAGTTATTTTTA	1860
Dd	1801	TTTTCTGAAAGCTAGAGAATTACAACCTTTGGTGATAAACCAAAAGGAGTCAGTTATTTTTA	1860
QY	1861	TTCATATATTTTAGCATATTCGAACCTAATTTCTAAGAAAATTTAGTTATAACTCTATGTAG	1920
Dd	1861	TTTCATATATTTTAGCATATTCGAACCTAATTTCTAAGAAAATTTAGTTATAACTCTATGTAG	1920
QY	1921	TTATAGAAAGTCAATATGCAGTTATTTCTATCAGTCGCACAATTCITGAGTCTCTGATAACC	1980
Dd	1921	TTATAGAAAGTCAATATGCAGTTATTTCTATCAGTCGCACAATTCITGAGTCTCTGATAACC	1980
QY	1981	TACCTATTGGGGTTAGGAGAAAAGACTAGACAATTACTATGTGGTCATTTCTCTACAACAT	2040
Dd	1981	TACCTATTGGGGTTAGGAGAAAAGACTAGACAATTACTATGTGGTCATTTCTCTACAACAT	2040
QY	2041	ATGTTAGCACGGCAAAGAACCCTTCAAATTTGAAGACTGAGATTTTCTCTGATATATATGGGTT	2100
Dd	2041	ATGTTAGCACGGCAAAGAACCCTTCAAATTTGAAGACTGAGATTTTCTCTGATATATATGGGTT	2100
QY	2101	TTGTAAAGATGGTTTTACACACTACAGATGCTCTATACTGTGAAAAAGTGTTTTCAATTCTG	2160
Dd	2101	TTGTAAAGATGGTTTTACACACTACAGATGCTCTATACTGTGAAAAAGTGTTTTCAATTCTG	2160
QY	2161	AAAAAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAGGTAGAGCTGTTCTTAAATT	2220
Dd	2161	AAAAAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAGGTAGAGCTGTTCTTAAATT	2220
QY	2221	TATTAATAAAAAAAAAAAAAA	2239
Dd	2221	TATTAATAAAAAAAAAAAAAA	2239

```

RESULT 3
US-09-854-133-441
; Sequence 441, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 441
; LENGTH: 5981
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-441

Query Match      83.4%;   Score 1867;   DB 9;   Length 5981;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1867;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps

QY      334  GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278  GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 337

QY      394  ACCATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGTGCTCCAGAACACCGGGCAGC 453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338  ACCATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGTGCTCCAGAACACCGGGCAGC 397

QY      454  GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACTATTTTGGAGCTTTTG 513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398  GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACTATTTTGGAGCTTTTG 457

QY      514  TCTTATGCTGAATTGGGAACAACTATARAAGAAATCTGGAGGTCTATTACACATATATTTTG 573

```

Db 1538 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGACAAGAAACCCAGG 1597
QY 1654 TGGTTTGAATAAATGTCAGAGAAAATAACCAACATTAACAAATAACTGGAAGTTGTA 1713
Db 1598 TGGTTTGAATAAATGTCAGAGAAAATAACCAACATTAACAAATAACTGGAAGTTGTA 1657
QY 1714 CCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCTTGGAATCTGCCCAAGGGA 1773
Db 1658 CCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCTTGGAATCTGCCCAAGGGA 1717
QY 1774 GACACAAAATAGGGATTTTACTTTTCTGAAAGTCTAGAGAAATTAACAATTTGGTG 1833
Db 1718 GACACAAAATAGGGATTTTACTTTTCTGAAAGTCTAGAGAAATTAACAATTTGGTG 1777
QY 1834 ATAAACAAAAGGAGTCAGTTATTTTATTTATTCATATATTTTAGCATATTTGCAACTAATTTCT 1893
Db 1778 ATAAACAAAAGGAGTCAGTTATTTTATTTATTCATATATTTTAGCATATTTGCAACTAATTTCT 1837
QY 1894 AAGAAATTTAGTTATTAATCTATGATAGTTATAGAAAGTGAATATGCAGTTATTTCTATGAG 1953
Db 1838 AAGAAATTTAGTTATTAATCTATGATAGTTATAGAAAGTGAATATGCAGTTATTTCTATGAG 1897
QY 1954 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAA 2013
Db 1898 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAA 1957
QY 2014 TTACTATGTGGTCAATCTCTACAACATATGTTAGCAGCGCAAGAACCTTCAAATTTGAAG 2073
Db 1958 TTACTATGTGGTCAATCTCTACAACATATGTTAGCAGCGCAAGAACCTTCAAATTTGAAG 2017
QY 2074 ACTGAGATTTTCTGTATATATGTTTGTAAAGATGTTTACACACTACAGATGTCT 2133
Db 2018 ACTGAGATTTTCTGTATATATGTTTGTAAAGATGTTTACACACTACAGATGTCT 2077
QY 2134 ATACTGTGAAAGTGTCTTCAATCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 2193
Db 2078 ATACTGTGAAAGTGTCTTCAATCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 2137
QY 2194 GAGAGAA 2200
Db 2138 GAGAGAA 2144

RESULT 4
US-09-738-973-441
; Sequence 441, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 441
; LENGTH: 5981
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-973-441

Query Match 83.4%; Score 1867; DB 10; Length 5981;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGSC 393
Db 278 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGSC 337
QY 394 ACCATATTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGAGC 453
Db 338 ACCATATTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGAGC 397
QY 454 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGCTCACTATTTGGAGCTTG 513
Db 398 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGCTCACTATTTGGAGCTTG 457
QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCACTACACATATATTTG 573
Db 458 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCACTACACATATATTTG 517
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGAACTCCTCATATATACGC 633
Db 518 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGAACTCCTCATATATACGC 577
QY 634 CTTGACGACTACTGCTGTGATATCCCTGGCATTTGGACGTACATTTGGAACCAATTTT 693
Db 578 CTTGACGACTACTGCTGTGATATCCCTGGCATTTGGACGTACATTTGGAACCAATTTT 637
QY 694 ATTCAATGTGAAATCCCTGAACCTTGGATCAAGCTACATTTGGGTGAACTCCTCATATATACGC 753
Db 638 ATTCAATGTGAAATCCCTGAACCTTGGATCAAGCTACATTTGGGTGAACTCCTCATATATACGC 697
QY 754 GTGATGGTCTTAATAGCATGAGTGTGAGCTGGAGGCGGATCCAGATTTTCTTAACC 813
Db 698 GTGATGGTCTTAATAGCATGAGTGTGAGCTGGAGGCGGATCCAGATTTTCTTAACC 757
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATAA 873
Db 758 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATAA 817
QY 874 GGTCAAAACGCAGAACCTTAAAGACGCCCTTTTCAAGAGAGATTCAAGTATTACGCGGTTG 933
Db 818 GGTCAAAACGCAGAACCTTAAAGACGCCCTTTTCAAGAGAGATTCAAGTATTACGCGGTTG 877
QY 934 CCACTGGCTTTTATATGGAATGTATGATATGCTGGCTGGTGTTCACCTCAACTTTGTT 993
Db 878 CCACTGGCTTTTATATGGAATGTATGATATGCTGGCTGGTGTTCACCTCAACTTTGTT 937
QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCAATTCCTTGTCAATATGATATATCCATGGCC 1053
Db 938 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCAATTCCTTGTCAATATGATATATCCATGGCC 997
QY 1054 ATTGTACCAATTTGGCTATGTGCTGACAAATCTGGCTACTTTACGACCAATTAATGCTGAG 1113
Db 998 ATTGTACCAATTTGGCTATGTGCTGACAAATCTGGCTACTTTACGACCAATTAATGCTGAG 1057
QY 1114 GAGCTGCTGCTTCAAAATGCAAGTGGCAGTACCTTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 1058 GAGCTGCTGCTTCAAAATGCAAGTGGCAGTACCTTTTCTGAGCGGCTACTGGGAAATTTTC 1117
QY 1174 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTG 1233
Db 1118 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTG 1177
QY 1234 TTTGCTGCTCCAGGTTATTCTATGTTGGTCTCGAGAGGGTCACTTCCAGAAATCCTC 1293
Db 1178 TTTGCTGCTCCAGGTTATTCTATGTTGGTCTCGAGAGGGTCACTTCCAGAAATCCTC 1237
QY 1294 TCCATGATTTCATGTCGCAAGCAGACACTCCTCTACCAGCTGTTATGTTTGCACCCCTTG 1353
Db 1238 TCCATGATTTCATGTCGCAAGCAGACACTCCTCTACCAGCTGTTATGTTTGCACCCCTTG 1297
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCTCAGTTTGGC 1413


```

1298 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTGAATTTCCCTCAGTTTGCC 1357
1414 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATATCTCGATACAAATGCCCCA 1473
1358 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATATCTCGATACAAATGCCCCA 1417
1474 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCCGAGCTTTGTTTTCCTTCACATGC 1533
1418 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCCGAGCTTTGTTTTCCTTCACATGC 1477
1534 CTCTTCATGGTTGCCCTTTCCCTCTATTCGGACCCCATTTAGTACAGGGATTGGCTTCGTC 1593
1478 CTCTTCATGGTTGCCCTTTCCCTCTATTCGGACCCCATTTAGTACAGGGATTGGCTTCGTC 1537
1594 ATCACTCTGACTGGAGTCCCTGCGTATATCTCTTTTATTATATGGGACAAGAAACCCAGG 1653
1538 ATCACTCTGACTGGAGTCCCTGCGTATATCTCTTTTATTATATGGGACAAGAAACCCAGG 1597
1654 TGGTTTAGAATAATGTGAGAGAAATAACAGAACATTTACAAATAACTGGAAGTTGTA 1713
1598 TGGTTTAGAATAATGTGAGAGAAATAACAGAACATTTACAAATAACTGGAAGTTGTA 1657
1714 CCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGA 1773
1658 CCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGA 1717
1774 GACACAAAATAGGGATTTTACTTTCATATATATTTAGCATATTCGAACTTAATTTCT 1833
1718 GACACAAAATAGGGATTTTACTTTCATATATATTTAGCATATTCGAACTTAATTTCT 1777
1834 ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTAGCATATTCGAACTTAATTTCT 1893
1778 ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTAGCATATTCGAACTTAATTTCT 1837
1894 AAGAAATTTAGTTATAAATCTATGTAGTTATAGAAAGTGAATATGCACTTATCTATGAG 1953
1838 AAGAAATTTAGTTATAAATCTATGTAGTTATAGAAAGTGAATATGCACTTATCTATGAG 1897
1954 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAA 2013
1898 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAA 1957
2014 TTAATATGTTGTCATCTCTACAAATATGTTAGCAGCGCAAGAACCTTCAAAATGAAG 2073
1958 TTAATATGTTGTCATCTCTACAAATATGTTAGCAGCGCAAGAACCTTCAAAATGAAG 2017
2074 ACTGAGATTTTCTGTATATATGGTTTGTAAAGATGGTTTACACACTACAGATGTCT 2133
2018 ACTGAGATTTTCTGTATATATGGTTTGTAAAGATGGTTTACACACTACAGATGTCT 2077
2134 ATACTGTGAAAAGTGTTCATTTCTGAAAAAAGCATACATCATGATTATGGCAAGAG 2193
2078 ATACTGTGAAAAGTGTTCATTTCTGAAAAAAGCATACATCATGATTATGGCAAGAG 2137
2194 GAGAGAA 2200
2138 GAGAGAA 2144
```

```

RESULT 5
US-10-163-866-31
; Sequence 31, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
```

```

; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 3144
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-163-866-31

Query Match      81.1%; Score 1816; DB 9; Length 3144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
Db 106 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 165
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCTCCAGAACACGGGCAGC 453
Db 166 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCTCCAGAACACGGGCAGC 225
QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTCTGGGGTCTCTGACATATTTGGAGCTTTG 513
Db 226 GTGGGCATGTCTCTGACCATCTGGACGGTGTCTGGGGTCTCTGACATATTTGGAGCTTTG 285
QY 514 TCTTATGTCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCACTTACACATATATTTTG 573
Db 286 TCTTATGTCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCACTTACACATATATTTTG 345
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAATACGC 633
Db 346 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAATACGC 405
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTTTT 693
Db 406 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTTTT 465
QY 694 ATTCAATGTGAAATCCCTGAACTTGGCATCAAGCTCATACAGCTGTGGGCACTAACTGTA 753
Db 466 ATTCAATGTGAAATCCCTGAACTTGGCATCAAGCTCATACAGCTGTGGGCACTAACTGTA 525
QY 754 GTGATGTCCTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACC 813
Db 526 GTGATGTCCTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACC 585
QY 814 TTTTGAAGCTCACAGCAATTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACCGGTTG 873
Db 586 TTTTGAAGCTCACAGCAATTTTCTGATAAATATAGTCCCTGGAGTTATGCAGCTAATAAAA 645
QY 874 GGTCAAAACGCAGAACTTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACCGGTTG 933
Db 646 GGTCAAAACGCAGAACTTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACCGGTTG 705
QY 934 CCACCTGCTTTTATTATGGAATGTATGATATGCTGGTGGTGTTCACCTCAACTTTGTT 993
Db 706 CCACCTGCTTTTATTATGGAATGTATGATATGCTGGTGGTGTTCACCTCAACTTTGTT 765
QY 994 ACTGAAGAGTAGAAAAACCCCTGAAAAAACCATTTCCCTTGAATATGTATATCCATGGCC 1053
Db 766 ACTGAAGAGTAGAAAAACCCCTGAAAAAACCATTTCCCTTGAATATGTATATCCATGGCC 825
QY 1054 ATTGTACCATTTGGCTATGTGTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 1113
Db 826 ATTGTACCATTTGGCTATGTGTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 885
QY 1114 GAGCTGCTGCTTTCAAATGCAGTGGAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 1114 GAGCTGCTGCTTTCAAATGCAGTGGAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTTC 1173
```


Db 886 GAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGGGCTACTGGGAAATTC 945

QY 1174 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTG 1233

Db 946 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTG 1005

QY 1234 TTTGCTGCTCCAGGTTATTCATGTTGCGTCTCGAGAGGCTCACCTTCCAGAAATCCTC 1293

Db 1006 TTTGCTGCTCCAGGTTATTCATGTTGCGTCTCGAGAGGCTCACCTTCCAGAAATCCTC 1065

QY 1294 TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTGGCACCCTTG 1353

Db 1066 TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTGGCACCCTTG 1125

QY 1354 ACAATGATAATGCTCTCTCTCGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGC 1413

Db 1126 ACAATGATAATGCTCTCTCTCGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGC 1185

QY 1414 AGGTGGCTTTTATTGGGCTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473

Db 1186 AGGTGGCTTTTATTGGGCTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1245

QY 1474 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTCAATCCAGCTTTGTTTCCCTCAGATGC 1533

Db 1246 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTCAATCCAGCTTTGTTTCCCTCAGATGC 1305

QY 1534 CTCTTCATGGTTGCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1593

Db 1306 CTCTTCATGGTTGCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1365

QY 1594 ATCACTCTGACTGGAGTCCCTCGGTATTATCTCTTTATATATGGGACAGAAACCCAGG 1653

Db 1366 ATCACTCTGACTGGAGTCCCTCGGTATTATCTCTTTATATATGGGACAGAAACCCAGG 1425

QY 1654 TGGTTTAGAATAATGTACAGAGAAATAACCCAGAACATTAACAAATAACTGGAAGTTGTA 1713

Db 1426 TGGTTTAGAATAATGTACAGAGAAATAACCCAGAACATTAACAAATAACTGGAAGTTGTA 1485

QY 1714 CCAGAAGAAGATAAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGA 1773

Db 1486 CCAGAAGAAGATAAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGA 1545

QY 1774 GACACAAAATAGGGATTTTACTTCAATTTTCTGAAAGTCTAGAGAATTACAACTTTGGTG 1833

Db 1546 GACACAAAATAGGGATTTTACTTCAATTTTCTGAAAGTCTAGAGAATTACAACTTTGGTG 1605

QY 1834 ATAAACAAAAGGAGTCAGTTATTTTATTTCATATATTTTAGCATATATTTAGCATATTTTCT 1893

Db 1606 ATAAACAAAAGGAGTCAGTTATTTTATTTCATATATTTTAGCATATATTTTCTGAACTAATTTCT 1665

QY 1894 AAGAAATTTAGTTATAACTCTATGATACCTACCTATTTAGGAGTGAATATGCACTTATCTATGAG 1953

Db 1666 AAGAAATTTAGTTATAACTCTATGATACCTACCTATTTAGGAGTGAATATGCACTTATCTATGAG 1725

QY 1954 TCGCACAAATTTCTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAA 2013

Db 1726 TCGCACAAATTTCTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAA 1785

QY 2014 TTACTATGTTGGTCATTTCTACAACATATGTTAGCACGGCAAGAACCTTCAAATTTGAAG 2073

Db 1786 TTACTATGTTGGTCATTTCTACAACATATGTTAGCACGGCAAGAACCTTCAAATTTGAAG 1845

QY 2074 ACTGAGATTTTCTGTATATATGGTTTGTAAAGATGGTTTACACACTACAGATGTCT 2133

Db 1846 ACTGAGATTTTCTGTATATATGGTTTGTAAAGATGGTTTACACACTACAGATGTCT 1905

QY 2134 ATACTGTGAAAGTGTTTTCAATTTCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 2193

Db 1906 ATACTGTGAAAGTGTTTTCAATTTCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 1965

QY 2194 GAGAGAA 2200

Db 1966 GAGAGAA 1972

RESULT 6

US-10-163-866-29

; Sequence 29, Application US/10163866

; Publication No. US20030027188A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SL7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/10/163,866

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: US 60/357,600

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 29

; LENGTH: 2482

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-163-866-29

Query Match 76.9%; Score 1722; DB 9; Length 2482;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1772; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGSC 393

Db 337 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGSC 396

QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 453

Db 397 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 456

QY 454 GTGGCATGCTCTGTGACCATCTGACGCGGTGTGTGGGGTCTCTGCTCACTATTTGGAGCTTG 513

Db 457 GTGGCATGCTCTGTGACCATCTGACGCGGTGTGTGGGGTCTCTGCTCACTATTTGGAGCTTG 516

QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCACTACACATATATTTG 573

Db 517 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCACTACACATATATTTG 576

QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATACGC 633

Db 577 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATACGC 636

QY 634 CTTGCAGTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTGGAACCAATTTT 693

Db 637 CTTGCAGTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTGGAACCAATTTT 696

QY 694 ATTCAATGTGAAATCCCTGAACTTGGATCAAGCTCAAGCTGTGGGCATAAATGTA 753

Db 697 ATTCAATGTGAAATCCCTGAACTTGGATCAAGCTCAAGCTGTGGGCATAAATGTA 756

QY 754 GTGATGGTCTCAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTCTTAACC 813

Db 757 GTGATGGTCTCAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTCTTAACC 816

QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGTAATAA 873

Db 817 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGTAATAA 876

QY 874 GGTCAAAACGCAGAACTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAAGTATTACGGGTTG 933

Db 877 GGTCAAAACGCAGAACTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAAGTATTACGGGTTG 936

QY 934 CCACTGGCTTTTATATGAATGATGATGATGCTGGCTGGTTTACCTCAACTTTGTT 993
Db 937 CCACTGGCTTTTATATGAATGATGATGATGCTGGCTGGTTTACCTCAACTTTGTT 996
QY 994 ACTGAAGAAGTAGAAACCCCTGAAAAACCATTCCTCCCTTGCAATATGATATATCCATGGCC 1053
Db 997 ACTGAAGAAGTAGAAACCCCTGAAAAACCATTCCTCCCTTGCAATATGATATATCCATGGCC 1056
QY 1054 ATTGTCAACCAATTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATAATGCTGAG 1113
Db 1057 ATTGTCAACCAATTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATAATGCTGAG 1116
QY 1114 GAGCTGCTGCTTTCAAAATGCGAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 1117 GAGCTGCTGCTTTCAAAATGCGAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1176
QY 1174 TCATTAGCAGTTCCGATCTTTGTTGCTGCTCCCTCTCCCTGCTCCATGAGCGTGGTGTG 1233
Db 1177 TCATTAGCAGTTCCGATCTTTGTTGCTGCTCCCTCTCCCTGCTCCATGAGCGTGGTGTG 1236
QY 1234 TTTGCTGCTCCAGGTTATTTCTATGTTGCTGCTCGAGAGGGTCACTTCCAGAAATCCTC 1293
Db 1237 TTTGCTGCTCCAGGTTATTTCTATGTTGCTGCTCGAGAGGGTCACTTCCAGAAATCCTC 1296
QY 1294 TCCATGATTCATGTCGCGCAAGCACACTCTCTACCAGCTGTTATTGTTTGGACCCCTTG 1353
Db 1297 TCCATGATTCATGTCGCGCAAGCACACTCTCTACCAGCTGTTATTGTTTGGACCCCTTG 1356
QY 1354 ACAATGATGATGCTCTCTCTGAGAGACCTCGACAGCTTTTGAATTTCCCTCAGTTTGCC 1413
Db 1357 ACAATGATGATGCTCTCTCTGAGAGACCTCGACAGCTTTTGAATTTCCCTCAGTTTGCC 1416
QY 1414 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 1417 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1476
QY 1474 GATATGATGCTCTCTCTCAAGGTGGCAGTGTTCATCCAGCTTTGTTTCCCTTCATGTC 1533
Db 1477 GATATGATGCTCTCTCTCAAGGTGGCAGTGTTCATCCAGCTTTGTTTCCCTTCATGTC 1536
QY 1534 CTCTTCATGCTGCCCTTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTC 1593
Db 1537 CTCTTCATGCTGCCCTTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTC 1596
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATTATATGGGACAAGAACCCAGG 1653
Db 1597 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATTATATGGGACAAGAACCCAGG 1656
QY 1654 TGGTTTAGAATAATGTGAGAGAAATAACAGAACATTACAAATAATCTGGAAGTTGTA 1713
Db 1657 TGGTTTAGAATAATGTGAGAGAAATAACAGAACATTACAAATAATCTGGAAGTTGTA 1716
QY 1714 CCAGAAGAAGATAAGTTATGAACATAATGGACTTGGCAATCTGCCCAAGGGA 1773
Db 1717 CCAGAAGAAGATAAGTTATGAACATAATGGACTTGGCAATCTGCCCAAGGGA 1776
QY 1774 GACACAAAATAGGGATTTTACTTCATTTTCTGAAAGTCTAGAGAATTACAACCTTTGGTG 1833
Db 1777 GACACAAAATAGGGATTTTACTTCATTTTCTGAAAGTCTAGAGAATTACAACCTTTGGTG 1836
QY 1834 ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCT 1893
Db 1837 ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCT 1896
QY 1894 AAGAAATTTAGTTAACTCTATGTAGTTATAGAAAGTGAATATGCAAGTTATTTCTATGAG 1953
Db 1897 AAGAAATTTAGTTAACTCTATGTAGTTATAGAAAGTGAATATGCAAGTTATTTCTATGAG 1956
QY 1954 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTGGGGTTAGGAGAAAGACTAGACAA 2013
Db 1957 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTGGGGTTAGGAGAAAGACTAGACAA 2016

QY 2014 TTACTATGTGGTCAATCTCTACACATATGTTAGCAGGCAAAAGAACCTTCAAATGAAG 2073
Db 2017 TTACTATGTGGTCAATCTCTACACATATGTTAGCAGGCAAAAGAACCTTCAAATGAAG 2076
QY 2074 ACTGAGATTTTCTGTATATATGGGTTTGTAA 2106
Db 2077 ACTGAGATTTTCTGTATATATGGGTTTGTAA 2109

RESULT 7
US-10-163-866-30
; Sequence 30, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 1861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-30

Query Match 64.9%; Score 1454; DB 9; Length 1861;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGSC 393
Db 341 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGSC 400
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTTAAGGGCGTCTCCAGAACACGGGCAGC 453
Db 401 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTTAAGGGCGTCTCCAGAACACGGGCAGC 460
QY 454 GTGGGCATGCTCTGACCATCTGGACGCTGTGGGCTCTGTCTGAGCTTTGAGCTTTG 513
Db 461 GTGGGCATGCTCTGACCATCTGGACGCTGTGGGCTCTGTCTGAGCTTTGAGCTTTG 520
QY 514 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTACACATATATTTG 573
Db 521 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTACACATATATTTG 580
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATACGC 633
Db 581 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATACGC 640
QY 634 CCTGCAGTACTGCTGTGATATCCCTGGCAATCCCTGGCAATTTGGACCTACTTTT 693
Db 641 CCTGCAGTACTGCTGTGATATCCCTGGCAATCCCTGGCAATTTGGACCTACTTTT 700
QY 694 ATTCAATGTGAATCCCTGAACTTGGATCGGATCAAGCTCATACAGTGTGGGCATAACTGTA 753
Db 701 ATTCAATGTGAATCCCTGAACTTGGATCGGATCAAGCTCATACAGTGTGGGCATAACTGTA 760
QY 754 GTGATGGTCCCTAAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTCTTAACC 813
Db 761 GTGATGGTCCCTAAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTCTTAACC 820

QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTAATAGTCCCTGGAGTTATGCAGCTAATATAA 873
Db 821 TTTTGAAGCTCACAGCAATTTCTGATAATTAATAGTCCCTGGAGTTATGCAGCTAATATAA 880
QY 874 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACGCGGTTG 933
Db 881 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACGCGGTTG 940
QY 934 CCACTGGCTTTTATATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 993
Db 941 CCACTGGCTTTTATATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 1000
QY 994 ACTGAAGAAGTAGAAACCCCTGAAAAACCCATTTCCCTTGCATATATGATATCCATGGCC 1053
Db 1001 ACTGAAGAAGTAGAAACCCCTGAAAAACCCATTTCCCTTGCATATATGATATCCATGGCC 1060
QY 1054 ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 1113
Db 1061 ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 1120
QY 1114 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGACGGGCTACTGGGAAATTC 1173
Db 1121 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGACGGGCTACTGGGAAATTC 1180
QY 1174 TCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTG 1233
Db 1181 TCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTG 1240
QY 1234 TTTGCTGTCTCAGGTTATTTCTATGTTGGCTCTCGAGAGGTCACCTTCCAGAAATCCTC 1293
Db 1241 TTTGCTGTCTCAGGTTATTTCTATGTTGGCTCTCGAGAGGTCACCTTCCAGAAATCCTC 1300
QY 1294 TCCATGATTATGTCGCAAGCACACTCCTCTACCAGCTGTTATTTGTTTGGCACCCTTTG 1353
Db 1301 TCCATGATTATGTCGCAAGCACACTCCTCTACCAGCTGTTATTTGTTTGGCACCCTTTG 1360
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGCTCTTTGAATTTCTCAGCTTTGCC 1413
Db 1361 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGCTCTTTGAATTTCTCAGCTTTGCC 1420
QY 1414 AGTGCTGCTTTTATTTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 1421 AGTGCTGCTTTTATTTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1480
QY 1474 GATATGATGCTGCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATCATGC 1533
Db 1481 GATATGATGCTGCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATCATGC 1540
QY 1534 CTCTTCATGCTTCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATGGCTTCGTC 1593
Db 1541 CTCTTCATGCTTCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATGGCTTCGTC 1600
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGAGGACAAACCCAGG 1653
Db 1601 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGAGGACAAACCCAGG 1660
QY 1654 TGGTTAGAAATAATGTCAGAGAAAATAACAGAACATTAACAATAATGGAAGTTGTA 1713
Db 1661 TGGTTAGAAATAATGTCAGAGAAAATAACAGAACATTAACAATAATGGAAGTTGTA 1720
QY 1714 CCAGAGAAGATAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGA 1773
Db 1721 CCAGAGAAGATAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGA 1780
QY 1774 GACACAAAATAGGGATTTTACTTCTTCAAAAGTCTAGAGAAATACAACCTTTGGTG 1833
Db 1781 GACACAAAATAGGGATTTTACTTCTTCAAAAGTCTAGAGAAATACAACCTTTGGTG 1840
QY 1834 ATAA 1838
Db 1841 ATAA 1845

RESULT 8
US-10-163-866-33
; Sequence 33, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-33

Query Match 63.4%; Score 1420; DB 9; Length 1542;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GAGAAAGTGCAGCTGAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
Db 114 GAGAAAGTGCAGCTGAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 173
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTCTCCAGAACACGGGCAGC 453
Db 174 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTCTCCAGAACACGGGCAGC 233
QY 454 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCTACATTTTGGAGCTTTG 513
Db 234 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCTACATTTTGGAGCTTTG 293
QY 514 TCTTATGCTGAATTTGGAAACAACATAAAGAAATCTGGAGGTCATACACATATATTTTG 573
Db 294 TCTTATGCTGAATTTGGAAACAACATAAAGAAATCTGGAGGTCATACACATATATTTTG 353
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATAATACGC 633
Db 354 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATAATACGC 413
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 693
Db 414 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 473
QY 694 ATTCAATGTGAATCCCTGAACTTGGCATCAAGCTCATTTACAGTGTGGGCATAACTGTA 753
Db 474 ATTCAATGTGAATCCCTGAACTTGGCATCAAGCTCATTTACAGTGTGGGCATAACTGTA 533
QY 754 GTGATGGTCTTAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTTCTTAACC 813
Db 534 GTGATGGTCTTAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTTCTTAACC 593
QY 814 TTTTGCAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGCAGCTAATATAA 873
Db 594 TTTTGCAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGCAGCTAATATAA 653
QY 874 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACGCGGTTG 933
Db 654 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACGCGGTTG 713
QY 934 CCACTGGCTTTTATTATGGAATGATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 993

Db 1081 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCTCTGCTTGGCTCCATGAACGGTGGTGTG 1140
QY 1234 TTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGAGAGGGTCAACCTTCCAGAAATCCTC 1293
Db 1141 TTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGAGAGGGTCAACCTTCCAGAAATCCTC 1200
QY 1294 TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATGTTTTCACCCCTTTG 1353
Db 1201 TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATGTTTTCACCCCTTTG 1260
QY 1354 ACAATGATTAATGCTCTCTCTGAGACCTCGACAGCTCTTTTGAATTTCCCTCAGTTTGGC 1413
Db 1261 ACAATGATTAATGCTCTCTCTGAGACCTCGACAGCTCTTTTGAATTTCCCTCAGTTTGGC 1320
QY 1414 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCA 1473
Db 1321 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCA 1380
QY 1474 GATATGATCGTCTCTTTCAAGGTGCCACTGTTTCATCCCGAGCTTTGTTTTCCTTCACATGC 1533
Db 1381 GATATGATCGTCTCTTTCAAGGTGCCACTGTTTCATCCCGAGCTTTGTTTTCCTTCACATGC 1440
QY 1534 CTCCTCATGTTGCCCTTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1593
Db 1441 CTCCTCATGTTGCCCTTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1500
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGAACAAGAACCCAGG 1653
Db 1501 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGAACAAGAACCCAGG 1560
QY 1654 TGGTTAGATAAATGTCAG 1672
Db 1561 TGGTTAGATAAATGTCAG 1579

RESULT 10
US-10-163-866-52
; Sequence 52, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-52

Query Match 58.0%; Score 1299; DB 9; Length 1528;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 334 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTCATTTGGC 393
Db 128 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTCATTTGGC 187
QY 394 ACCATCATGGAGCAGGAATCTTCATCTCTCCTAAGGCGTGTCTCCAGAACACGGGCAGC 453

Db 188 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTAAGGGCGTCTCCAGAACACGGGCAGC 247
QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACTATTTGGAGCTTTG 513
Db 248 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACTATTTGGAGCTTTG 307
QY 514 TCTTATGCTGAATTTGGGAACAACACTATAAAGAAATCTGGAGGTCATTACACATATATTTTG 573
Db 308 TCTTATGCTGAATTTGGGAACAACACTATAAAGAAATCTGGAGGTCATTACACATATATTTTG 367
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 633
Db 368 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 427
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTTGGAACCATTTTGT 693
Db 428 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTTGGAACCATTTTGT 487
QY 694 ATTCAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTTACAGCTGTGGGCATAACTGTA 753
Db 488 ATTCAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTTACAGCTGTGGGCATAACTGTA 547
QY 754 GTGATGGTCCCTAAATAGCATGAGTGTGATTAATATAGTCCCTGGAGTATGTAATCTTAACC 813
Db 548 GTGATGGTCCCTAAATAGCATGAGTGTGATTAATATAGTCCCTGGAGTATGTAATCTTAACC 607
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGTAATCTTAACC 873
Db 608 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGTAATCTTAACC 667
QY 874 GGTCAAAACGAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTG 933
Db 668 GGTCAAAACGAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTG 727
QY 934 CCACCTGGCTTTTATATGGAATGTATGATATGCTGCTGGTGGTCTTACCTCAACTTTGT 993
Db 728 CCACCTGGCTTTTATATGGAATGTATGATATGCTGCTGGTGGTCTTACCTCAACTTTGT 787
QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTTCCCTTGCATATGTAATCTTAACC 1053
Db 788 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTTCCCTTGCATATGTAATCTTAACC 847
QY 1054 ATTGTACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 1113
Db 848 ATTGTACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 907
QY 1114 GAGCTGCTGCTTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 908 GAGCTGCTGCTTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 967
QY 1174 TCATTAGCAGTTCGGATCTTTGTTGCCCTCTCCGCTTTGGCTCCATGAACGGTGGTGTG 1233
Db 968 TCATTAGCAGTTCGGATCTTTGTTGCCCTCTCCGCTTTGGCTCCATGAACGGTGGTGTG 1027
QY 1234 TTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGAGAGGGTCAACCTTCCAGAAATCCTC 1293
Db 1028 TTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGAGAGGGTCAACCTTCCAGAAATCCTC 1087
QY 1294 TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATGTTTTCACCCCTTTG 1353
Db 1088 TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATGTTTTCACCCCTTTG 1147
QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCTCTCAGTTTGGC 1413
Db 1148 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCTCTCAGTTTGGC 1207
QY 1414 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCA 1473
Db 1208 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCA 1267
QY 1474 GATATGATCGTCTCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCACATGC 1533
Db 1268 GATATGATCGTCTCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCACATGC 1327

```
QY 1534 CTCCTCATGGTTGCCCTTTCCCTCTATTTCGGACCCCAATTTAGTACAGGGATTGGCTTCGTC 1593
Db 1328 CTCCTCATGGTTGCCCTTTCCCTCTATTTCGGACCCCAATTTAGTACAGGGATTGGCTTCGTC 1387

QY 1594 ATCACTCTGACTGGAGTCCCTCGGTATTATCTCTTTTATTATATGGGACAAGAAACCCAGG 1653
Db 1388 ATCACTCTGACTGGAGTCCCTCGGTATTATCTCTTTTATTATATGGGACAAGAAACCCAGG 1447

QY 1654 TGGTTTGAATAATGTGAGAGAAATAACCAAGAACATTACAAATAATACTGGAAGTTGTA 1713
Db 1448 TGGTTTGAATAATGTGAGAGAAATAACCAAGAACATTACAAATAATACTGGAAGTTGTA 1507

QY 1714 CCAGAAGAAGATAAGTTATGA 1734
Db 1508 CCAGAAGAAGATAAGTTATGA 1528

RESULT 11
US-10-163-866-53
; Sequence 53, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-53

Query Match 38.8%; Score 868; DB 9; Length 1268;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 128 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGC 187

QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTTAAGGGCGTGTCTCCAGAACACACGGGCAGC 453
Db 188 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTTAAGGGCGTGTCTCCAGAACACACGGGCAGC 247

QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTGCTCACTATTGGAGCTTTG 513
Db 248 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTGCTCACTATTGGAGCTTTG 307

QY 514 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTACACATATATTG 573
Db 308 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTACACATATATTG 367

QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAACGC 633
Db 368 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAACGC 427

QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATTCTGGAAACCATTTTTT 693
Db 428 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATTCTGGAAACCATTTTTT 487
```

```
QY 694 ATTCATATGTGAAATCCCTGAACCTTGGATCAAGCTCAATACAGCTGTGGGCATCAACTGTA 753
Db 488 ATTCATATGTGAAATCCCTGAACCTTGGATCAAGCTCAATACAGCTGTGGGCATCAACTGTA 547

QY 754 GTGATGGTCTCTAAATAGCATGAGTGTGATGAGTGGAGCGCCCGGATCCAGATTCTTAAACC 813
Db 548 GTGATGGTCTCTAAATAGCATGAGTGTGATGAGTGGAGCGCCCGGATCCAGATTCTTAAACC 607

QY 814 TTTTGCAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGCAGCTAATTTAAA 873
Db 608 TTTTGCAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGCAGCTAATTTAAA 667

QY 874 GGTCAAAACGAGAACTTTTAAAGACGCTTTTTCAGGAAGAGATCAAGTATTACGGGTTG 933
Db 668 GGTCAAAACGAGAACTTTTAAAGACGCTTTTTCAGGAAGAGATCAAGTATTACGGGTTG 727

QY 934 CCACTGGCTTTTATTATGGAATGATGATATGCTGGCTGGTGTACCTCAACTTTGTT 993
Db 728 CCACTGGCTTTTATTATGGAATGATGATATGCTGGCTGGTGTACCTCAACTTTGTT 787

QY 994 ACTGAAGAAGTAGAAAAACCCCTGAAAAAACCATTTCCCTTGCAATATGATATCCATGGCC 1053
Db 788 ACTGAAGAAGTAGAAAAACCCCTGAAAAAACCATTTCCCTTGCAATATGATATCCATGGCC 847

QY 1054 ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113
Db 848 ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 907

QY 1114 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTT 1173
Db 908 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTT 967

QY 1174 TCATTAGCAGTTCCGATCTTTGTTGCC 1201
Db 968 TCATTAGCAGTTCCGATCTTTGTTGCC 995

RESULT 12
US-10-163-866-32
; Sequence 32, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-32

Query Match 18.5%; Score 415; DB 9; Length 520;
Best Local Similarity 100.0%; Pred. No. 1.2e-201;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 106 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGC 165
```

Query Match	12.6%;	Score 281;	DB 9;	Length 337;
Best Local Similarity	100.0%;	Pred. No. 3.8e-133;		
Matches 281;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAC 60			
Db				
4	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAC 63			
QY 61	AGAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGGAAATTCACGGCCAATTGTGG 120			
Db				
64	AGAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGGAAATTCACGGCCAATTGTGG 123			
QY 121	CATAGATTTTATCATATTTCTGGATTTTTTGGATCTTTTGTCTCATCACTGGATTCA 180			
Db				
124	CATAGATTTTATCATATTTCTGGATTTTTTGGATCTTTTGTCTCATCACTGGATTCA 183			
QY 181	GGAAAGCCTGTTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAAATGTTAACGGG 240			
Db				
184	GGAAAGCCTGTTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAAATGTTAACGGG 243			
QY 241	AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCAGGA 281			
Db				
244	AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCAGGA 284			

```

RESULT 15
US-10-046-935-1307
; Sequence 1307, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 1307
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 9, 19, 461, 497, 500, 502
; OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1307

Query Match 4.8%; Score 107; DB 9; Length 572;
Best Local Similarity 100.0%; Pred. No. 3.3e-44;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTGGC 393
|||||
Db 354 GAGAAAGTGCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTGGC 413
|||||
QY 394 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGTGCTCCA 440
|||||
Db 414 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGTGCTCCA 460
|||||

Search completed: April 16, 2003, 16:06:46
Job time : 236 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: April 16, 2003, 14:10:10 ; Search time 71 Seconds
(without alignments)
8404.173 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_090917_20278/app_query.fasta_1.2375
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptn -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170_eCGN_1_1_79_erunat_04042003_090917_20278 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2519	61.1	521	22	ABB11957	Human cystine/Glu
2	2168	52.6	424	22	AAAM23914	Human EST encoded
3	1210.5	29.4	507	21	AAAY82488	Human L-type amino
4	1204	29.2	512	21	AAAY82489	Rat L-type amino a
5	1204	29.2	512	21	AAAY58426	Rat neutral amino
6	1140	27.7	507	22	ABB12114	Human 4F2 light ch
7	1131.5	27.5	511	20	AAAY41159	Human amino acid p
8	1131.5	27.5	511	20	AAAY27076	Human amino acid p
9	1131.5	27.5	511	21	AAAY87630	Human monp-2 prote
10	1113.5	27.0	505	22	ABB58292	Drosophila melanog
11	1108	26.9	533	22	AAAB72393	Rat L-type amino a
12	1106	26.8	499	22	ABB71209	Drosophila melanog
13	1101.5	26.7	535	21	AAAB43136	Human ORFX ORF2900
14	1101.5	26.7	535	21	AAAY71061	Human membrane tra
15	1101.5	26.7	535	22	AAAB72395	Human L-type amino
16	1091	26.5	517	22	ABB62049	Drosophila melanog
17	1083	26.3	564	22	ABB11788	Human amino acid t
18	1080.5	26.2	500	22	ABB62891	Drosophila melanog
19	1075.5	26.1	487	22	AAAB88569	Human hydrophobic
20	1051.5	25.5	530	22	AAG63803	Amino acid sequenc
21	1039	25.2	523	22	AAG63804	Amino acid sequenc
22	938.5	22.8	445	22	ABB64405	Drosophila melanog
23	741.5	18.0	352	21	AAAY91576	Human secreted pro
24	732	17.8	414	21	AAAB54295	Human secreted c
25	730	17.7	370	21	AAAY91451	Human secreted pro
26	714	17.3	386	22	ABG12125	Novel human diagno
27	603	14.6	377	21	AAAB10289	Human El6H protein
28	556.5	13.5	389	22	AAAB88577	Human hydrophobic
29	524.5	12.7	245	19	AAAW80803	Human tumour-assoc
30	475	11.5	194	23	ABB90268	Human polypeptide
31	415.5	10.1	222	23	ABB90216	Human polypeptide
32	406.5	9.9	179	22	ABB37681	Peptide #5187 enco
33	406.5	9.9	179	22	AAAM58312	Human brain expres
34	406.5	9.9	179	22	AAAM70786	Human bone marrow
35	406.5	9.9	179	22	AAAM18624	Peptide #5058 enco
36	406.5	9.9	179	22	AAAM31084	Peptide #5121 enco
37	406.5	9.9	179	23	ABG40523	Human peptide enco
38	331	8.0	463	23	ABB48552	Listeria monocytog
39	323	7.8	462	22	AAG98875	E. coli growth and
40	322	7.8	456	22	AAU33436	Enterococcus faeca
41	322	7.8	463	22	AAU35069	Enterococcus faeca
42	312.5	7.6	486	23	ABP29952	Streptococcus poly
43	307.5	7.5	583	21	AAG22206	Arabidopsis thalia
44	307.5	7.5	595	21	AAG22205	Arabidopsis thalia
45	305.5	7.4	538	21	AAG22207	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ABB11957

ID ABB11957 standard; peptide; 521 AA.

XX ABB11957;

XX 11-JAN-2002 (first entry)

XX Human cystine/Glu transporter homologue, SEQ ID NO:2327.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;

KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR N-PSDB; ABA09201.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer .
XX
PS Claim 20; Page 286; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 521 AA;

Alignment Scores:
Pred. No.: 3.36e-276 Length: 521
Score: 2519.00 Matches: 498
Percent Similarity: 96.51% Conservative: 0

Best Local Similarity: 96.51% Mismatches: 0	
Query Match: 61.14% Indels: 18	
DB: 22 Gaps: 1	
US-09-667-170A-440 (1-2239) x ABB11957 (1-521)	
QY 184	AAGCCTGTTGTCTCCACCATCTCCAAAGAGAGGTTACCTGCAGGGAAATGTTAACGGGAGG 243
Db 24	LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlyAsnValAsnGlyArg 43
QY 244	CTGCCTTCCCTGGGCAACAAAGAGAGCCACCTGGGAGGACGCGCTTTTCAGGAAGAGACGCC 303
Db 44	LeuProSerLeuGlyAsnLysGluProProGlyGln----- 55
QY 304	TTTTTCAGGAAGAGACGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACT 363
Db 56	-----GluLysValGlnLeuLysArgLysValThr 65
QY 364	TTACTGAGGGAGTCTCCATTATATATGTCACCATCATTTGGAGCAGGAATCTTTCATCTCT 423
Db 66	LeuLeuArgGlyValSerIleIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 85
QY 424	CCTAAGGGCGTCTCCAGAACACACGGGAGCGCTGGGCATGTCTCTGACCATCTGGACGGTG 483
Db 86	ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal 105
QY 484	TGTGGGCTCCTGTCACTATTGAGGCTTTGTCTTATGCTGAATGGGAACAACATATAAG 543
Db 106	CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrThrIleLys 125
QY 544	AAATCTGGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTGTFA 603
Db 126	LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal 145
QY 604	CGAGTCTGGGTGGAACCTCCTCATATACGACCCCTGCAGCTACTGTGTGATATCCCTGGCA 663
Db 146	ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla 165
QY 664	TTTGGACGCTACATTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTTGGGATC 723
Db 166	PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 185
QY 724	AAGCTCATTACAGCTGTGGGCATAACTGTAGTGTGATGGTCTCTAAATAGCATGAGTGTGAG 783
Db 186	LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 205
QY 784	TGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGCAAGCTCACAGCAATTTCTGATAAT 843
Db 206	TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIleIle 225
QY 844	ATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCGCAGAACTTTAAAGACGCCCTT 903
Db 226	IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 245
QY 904	TCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATGCA 963
Db 246	SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 265
QY 964	TATGCTGGCTGGTTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAAAC 1023
Db 266	TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 285
QY 1024	ATTCCTTTCGAATATGATATATCCATGGCCATTGTCAACCATGGCTATGTGTGACAAAT 1083
Db 286	IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 305
QY 1084	GTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCAGTG 1143
Db 306	ValAlaTyrPheThrThrIleAsnAlaGluGluLeuLeuLeuSerAsnAlaValAlaVal 325
QY 1144	ACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTTCGCGATCTTTGTTGCCCTC 1203
Db 326	ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 345

QY 1204 TCCTGCTTTGGCTCCATGAACGGTGGTGTGTGCTGTCTCCAGGTTATTCTATGTTGCG 1263
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1264 TCTCGAGAGGGTCCACCTTCCAGAAATCCCTCTCCATGATTCATGTCGCGCAAGCACACTCCT 1323
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1324 CTACCCAGCTGTTATGTTTTCACCCCTTGACAATGATAATGCTCTCTCTGGAGACCTC 1383
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1384 GACAGTCTTTTGAATTTCTCAGTTTTCGCGAGGTGGCTTTTATGTTGGCTGGCAGTTGCT 1443
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1444 GGGCTGATTATCTTCGATACAAAATGCCAGATATGCATCGTCTTCAAGGTGCCACTG 1503
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1504 TTCATCCAGCTTGTGTTTTCCTTCACATGCTCTCTTCATGTTGCTTCCCTCTCTATTCG 1563
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1564 GACCCATTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTAT 1623
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1624 CTCTTTTATATATGGGACAAAGAACCCAGGTGGTGTAGATAATGTACAGAGAAATAAAC 1683
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1684 AGAACATTACAAATAATACTGGAAGTTGTACCAGAGAAAGATAAGTTA 1731
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 2
AAM23914
ID AAM23914 standard; Protein; 424 AA.
XX
AC AAM23914;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST encoded protein SEQ ID NO: 1439.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
XX biodiversity; gene therapy; nutrition.
OS Homo sapiens.
XX
PN WO200154477-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US03687.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
DR WPI; 2001-476164/51.
DR N-PSDB; AAH98573.
XX

PT Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX
PS Claim 20; Page 1001-1002; 1275pp; English.
XX
CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.
XX
SQ Sequence 424 AA;

Alignment Scores:
Pred. No.: 2.08e-236 Length: 424
Score: 2168.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.62% Indels: 0
DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAM23914 (1-424)

QY 460 ATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTTGGAGCTTGTCTTAT 519
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeuSerTyr 20
QY 520 GCTGAATTGGGAACAACATATAAGAAATCTGGAGGTCACTACACATATATTTTGAAGTC 579
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21 AlaGluLeuGlyThrIleLysLysSerGlyGlyHistyrThrIleLeuGluVal 40
QY 580 TTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATACGCCCTGCA 639
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
41 PheGlyProLeuProAlaPheValArgValTrpValGluLeuLeuIleArgProAla 60
QY 640 GCTACTGCTGTGATATCCCTGGCATTGTCAGCTTACAGCTGTGGGATACCTTTTATTCAA 699
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 AlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPheIleGln 80
QY 700 TGTGAAATCCCTGAACCTTGGCATCAAGTCATACAGCTGTGGGATACCTGTAGTGATG 759
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
81 CysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrValValMet 100
QY 760 GTCCTAAATACCATGAGTGTCTGAGTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGC 819
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
101 ValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThrPheCys 120
QY 820 AAGCTCACAGCAATCTGATAATATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAA 879
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 LysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLysGlyGln 140
QY 880 ACGCAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGGGTGGCCACTG 939
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
141 ThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerSerIleThrArgLeuProLeu 160
QY 940 GCTTTTATTATGGAATGTATGATATGCTGGCTGGTGGTTTACCTCAACTTTGTTACTGAA 999
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
161 AlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheValThrGlu 180
QY 1000 GAAGTAGAAAACCCCTGAAAAACCATTTCCCTTGGCAATATGATATCCATGGCCATTGTC 1059
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 GluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAlaIleVal 200
QY 1060 ACCATTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTG 1119
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
201 ThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrIleAsnAlaGluLeu 220
QY 1120 CTGCTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTA 1179
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
221 LeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPheSerLeu 240

QY 1180 GCAGTCCGATCTTGTGTCCTCTCTGCTGGCTCCATGAACGGGTGTGTTGCT 1239
Db 241 AlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyValPheAla 260
QY 1240 GTCTCCAGGTTATTCTATGTTGCTCTCGAGAGGGTCCACTTCCAGAAATCCTCTCCATG 1299
Db 261 ValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluLeuSerMet 280
QY 1300 ATTATGTCGCGAAGCACACTCTCTACCAAGCTGTATTGTTTGCACCTTTGACAATG 1359
Db 281 IleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeuThrMet 300
QY 1360 ATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTGAATTTCTCCTCAGTTTGGCAGGTGG 1419
Db 301 IleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAlaArgTrp 320
QY 1420 CTTTATTATGGCTGGCTGGCTGGCTGATTTATCTTCGATACAAATGCCAGATATG 1479
Db 321 LeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysProAspMet 340
QY 1480 CATGTCCTTTCAAGGTGCCACTGTTTCATCCAGCTTGTGTTTCTCTCACATGCTCTTC 1539
Db 341 HisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCysLeuPhe 360
QY 1540 ATGGTTCCTCTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACT 1599
Db 361 MetValAlaLeuSerLeuTyrSerAspProPheSerThrGlyIleGlyPheValIleThr 380
QY 1600 CTGACTGGAGTCCCTGCGTATTATCTCTTTTATTATATGGGACAAAGAAACCCAGGTGTTT 1659
Db 381 LeuThrGlyValProAlaTyrTyrLeuPheIleIleTyrAspLysLysProArgTrpPhe 400
QY 1660 AGAATAATGTCAGAGAAATAACAGAACATTTACAAATAATACTGGAAGTTGTACCAGAA 1719
Db 401 ArgIleMetSerGluLysIleThrArgThrLeuGlnIleIleLeuGluValValProGlu 420
QY 1720 GAAGATAACTTA 1731
Db 421 GluAspLysLeu 424

RESULT 3
AA82488
ID AAY82488 standard; Protein; 507 AA.

AC AAY82488;
XX
DT 12-JUL-2000 (first entry)
DE. Human L-type amino acid transporter 1 protein sequence SEQ ID NO:2.
XX
KW L-type amino acid transporter 1; LAT1; 4F2hc; 4F2 heavy chain;
KW neutral amino acid transporter; tumour; cancer; proliferation;
KW cell membrane surface 4F2 molecule; anticancer.

OS Homo sapiens.
XX
PN WO200014228-A1.
XX
PD 16-MAR-2000.
XX
PF 03-SEP-1999; 99WO-JP04789.
XX
PR 03-SEP-1998; 98JP-0249993.
PR 02-SEP-1999; 99JP-0248546.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Endou H, Kanai Y;
XX
DR WPI; 2000-256979/22.
DR N-PSDB; AAA08372, AAA08393.
XX

PT Neutral amino acid transporter protein which conjugates with cell
PT membrane protein 4F2 and operates independently of sodium ions is
PT useful for screening potential cancer proliferation inhibitors -
XX
PS Claim 5; Page 133-136; 189pp; Japanese.
XX
CC The present invention describes a cell surface protein which is an
CC L-type amino acid transporter-1 (LAT1), which mediates the transport
CC of neutral amino acids, leucine, isoleucine, phenylalanine, methionine,
CC tyrosine, tryptophan, valine and histidine, into the cell independently
CC of sodium ions. The LAT1 protein conjugates with the cell membrane
CC surface molecule 4F2. Antisense DNA or RNA, antibodies, antibody
CC fragments, and peptide and non-peptide antagonists to LAT1 are useful
CC as anticancer agents. The present sequence represents human LAT1, which
CC is specifically claimed in the present invention.

XX Sequence 507 AA;

Alignment Scores:

Pred. No.: 1.13e-127 Length: 507
Score: 1210.50 Matches: 232
Percent Similarity: 66.19% Conservative: 95
Best Local Similarity: 46.96% Mismatches: 164
Query Match: 29.38% Indels: 3
DB: 21 Gaps: 2

US-09-667-170A-440 (1-2239) x AAY82488 (1-507)

QY 247 CCTTCCCTGGGCAACAAAGGAGCGCACCTGGGACGACGCTTTCAGGAAGAGACGCTTT 306
Db 14 ProAlaAlaGluGluLysGluAlaArgGluLysMetLeuAlaAlaLysSerAlaAsp 33
QY 307 TCAGGAAGAGAGACGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTA 366
Db 34 GlySerAlaProAlaGlyGluGly--GluGlyValThrLeuGlnArgAsnIleThrLeu 52
QY 367 CTGAGGGGAGTCTCCATTATCATTTGGCACCATTATGGAGCAGGAATCTTCTCTCTCT 426
Db 53 LeuAsnGlyValAlaIleIleValGlyThrIleIleGlySerGlyIlePheValThrPro 72
QY 427 AAGGGCTGCTCCAGAACACACGGGCGGCTGGGCTGCTCTGACCATCTGGACGGTGTGT 486
Db 73 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValTrpAlaAlaCys 92
QY 487 GGGTCTCTGTCACATATTGGAGCTTTGCTTATGCTGAATTTGGAACAACTATAAGAAA 546
Db 93 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleSerLys 112
QY 547 TCTGGAGGTCAATACACATATATTTTGAAGTCTTTGGTCCATTACCAGCTTTTGTACGA 606
Db 113 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCTGGGTGGAACCTCTCATATAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
Db 133 LeuTrpIleGluLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATCTGGAACCATTTTATTCATTAATGTGAATCCCTGAACCTTGGATCAAG 726
Db 153 AlaThrTyrLeuLysProLeuPheProThrCysProValProGluGluAlaAlaLys 172
QY 727 CTCATTACAGCTGTGGGCATAACTGTAGTGTGGTCTCTAAATAGCATGCTGCTGCTGCT 786
Db 173 LeuValAlaCysLeuCysValLeuLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla 192
QY 787 AGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATCTGATAATATA 846
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaAlaLysLeuLeuAlaLeuIleIle 212
QY 847 GTCCCTGGAGTTATGCGAGTAATTAAGGTCAACGACGAGAACTTTAAAGACGCTTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGAATGTAT 960

Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
QY 961 GCATATGCTGGCTGTTTACCTCAACTTGTACTGAAGAGTAGAAACCCCTGAAAA 1020
Db 253 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACCATTCCCTTGCATATGTATATCCATGGCCATTGTCCACCATGGCTATGTGCTGACA 1080
Db 273 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
QY 1081 AATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCA 1140
Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAACAGTCCGATCTTTGTTGCC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
QY 1201 CTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTTGTGCTCTCCAGGTTATCTATGTT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 352
QY 1261 GCGTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGAAGCACACT 1320
Db 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
QY 1321 CCTCTACCACTGTTATTGTTTGTGACCCCTTTGACAAATGATAATGCTCTCTCTGAGAC 1380
Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
QY 1381 CTCGACACTCTTTTGAATTTCTCCTCAGTTTGGCCAGTGGCTTTTATTGGGCTGGCAGTT 1440
Db 393 IlePheSerValIleAsnPheSerPhePheAsnTrpLeuCysValAlaLeuAlaIle 412
QY 1441 GCTGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGTCTCTTCAAGGTGCCA 1500
Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
QY 1501 CTGTTTCATCCAGCTTTGTTTCTTCCATGCTCTTCCATGCTTGGCTTCCCTCTCTAT 1560
Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452
QY 1561 TCGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGACGGAGTCCCTGCGTAT 1620
Db 453 LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr 472
QY 1621 TATCTCTTTATATATGGACAAAGAACCCAGGTGGTTTGAATAATATGTCAGAGAAATA 1680
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 492
QY 1681 ACCAGAACATTACAAATAATACTGGAAGTTGTACCAGAAGAA 1722
Db 493 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 506

RESULT 4
AAY82489
ID AAY82489 standard; protein; 512 AA.
XX
AC AAY82489;
XX
DT 12-JUL-2000 (first entry)
XX
DE Rat L-type amino acid transporter 1 protein sequence SEQ ID NO:4.
XX
KW L-type amino acid transporter 1; LAT1; 4F2hc; 4F2 heavy chain;
KW neutral amino acid transporter; tumour; cancer; proliferation;
KW cell membrane surface 4F2 molecule; anticancer.
XX
OS Rattus sp.
XX
PN W0200014228-A1.
XX

PD 16-MAR-2000.
XX
PF 03-SEP-1999; 99WO-JP04789.
XX
PR 03-SEP-1998; 98JP-0249993.
PR 02-SEP-1999; 99JP-0248546.
XX
(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Endou H, Kanai Y;
XX
WPI; 2000-256979/22.
DR N-PSDB; AAA08373.
XX
PT Neutral amino acid transporter protein which conjugates with cell membrane protein 4F2 and operates independently of sodium ions is useful for screening potential cancer proliferation inhibitors -
XX
PS Claim 5; Page 144-147; 189pp; Japanese.
XX
CC The present invention describes a cell surface protein which is an L-type amino acid transporter-1 (LAT1), which mediates the transport of neutral amino acids, leucine, isoleucine, phenylalanine, methionine, tyrosine, tryptophan, valine and histidine, into the cell independently of sodium ions. The LAT1 protein conjugates with the cell membrane surface molecule 4F2. Antisense DNA or RNA, antibodies, antibody fragments, and peptide and non-peptide antagonists to LAT1 are useful as anticancer agents. The present sequence represents rat LAT1, which is specifically claimed in the present invention.
XX
SQ Sequence 512 AA;

Alignment Scores:
Pred. No.: 6.22e-127 Length: 512
Score: 1204.00 Matches: 229
Percent Similarity: 67.57% Conservative: 96
Best Local Similarity: 47.61% Mismatches: 150
Query Match: 29.22% Indels: 6
DB: 21 Gaps: 2

US-09-667-170A-440 (1-2239) x AAY82489 (1-512)

QY 298 GACGCCTTTTCAGGAAGAGACGGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAA 357
Db 31 GluAlaArgArgGlyAspGlyAlaAspProGluGlyGluGlyValThrLeuGlnArgAsn 50
QY 358 GTCACCTTTACTGAGGGGAGTCTCCATATCATTTGGCACCATCATTTGGAGCAGGAATCTTC 417
Db 51 IleThrLeuIleAsnGlyValAlaIleValGlyThrIleIleGlySerGlyIlePhe 70
QY 418 ATCTCTCTTAAGGGCGTCTCCAGAACACGGCAGCGTGGGCATGTCTGTGACCATCTGG 477
Db 71 ValThrProThrGlyValLeuLysGluAlaGlySerProGlyLeuSerLeuValTrp 90
QY 478 ACGGTGTGTGGGCTCTGTCACTATTGGAGCTTTGTCTTATGTGAATTTGGGAACAAC 537
Db 91 AlaValCysGlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThr 110
QY 538 ATAAAGAAATCTGGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCT 597
Db 111 IleSerLysSerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAla 130
QY 598 TTTGTACGAGTCTGGGTGGAACCTCTCATATATACGCCCTGCAGCTACTGTGTATATCC 657
Db 131 PheLeuLysLeuTrpIleGluLeuLeuIleIleIleIleIleIleIleIleIleIleIle 150
QY 658 CTGGCATTTGGACGCTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCT 717
Db 151 LeuValPheAlaThrTyrLeuLeuLysProValPheProThrCysProValProGluGlu 170
QY 718 GCGATCAAGCTCATTAACAGCTGTGGGCATTAACCTAGTAGTGTGCTAAATAGCATGAGT 777
Db 171 AlaAlaLysLeuValAlaCysLeuValLeuLeuLeuThrAlaValAsnCysTyrSer 190

QY	718	CGGATCAAGCTCATTACAGCTGTGGGCATAACTGATGTAGTGGTCTCTAAATAGCATGAGT	777
Db	171	AlaAlaLysLeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTrpSer	190
QY	778	GTCAGCTGAGCGCCCGGATCCAGATTCTTAAACCTTTTGCAGAGCTCACAGCAATCTCTG	837
Db	191	ValLysAlaAlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAla	210
QY	838	ATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAA-----GGTCAAAACGCAG	885
Db	211	LeuIleLeuLeuGlyPheIleGlnMetGlyLysAspIleGlyGlnGlyAspAlaSer	230
QY	886	AACTTT-----AAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTG	939
Db	231	AsnLeuHisGlnLysLeuSerPheGluGlyThrAsnLeuAspValGlyAsnIleValLeu	250
QY	940	GCTTTTATTATGGAATGTATGCATATCTGGCTGGTGTATACCTCAACTTTGTTACTGAA	999
Db	251	AlaLeuTyrSerGlyLeuPheAlaTyrGlyGlyTyrAsnTyrLeuAsnPheValThrGlu	270
QY	1000	GAAGTAGAAAACCCCTGAAAAAACCATTCCTCTTGCAATATATATATCCATGGCCATTGTC	1059
Db	271	GluMetIleAsnProTyrArgAsnLeuProLeuAlaIleIleSerLeuProIleVal	290
QY	1060	ACCATTTGGCTATGTCTGCACAAATGTGGCCTACTTTACGACCATTAAATGCTGAGGAGCTG	1119
Db	291	ThrLeuValTyrValLeuThrAsnLeuAlaTyrPheThrLeuSerThrAsnGlnMet	310
QY	1120	CTGCTTCAAATGCAGTGGCAGTGCACCTTTTCTGAGCGGCTACTGGGAAATTTCTCAATTA	1179
Db	311	LeuThrSerGluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrp	330
QY	1180	GCAGTTCCGATCTTTGTTGGCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCT	1239
Db	331	IleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThr	350
QY	1240	GTCCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATG	1299
Db	351	SerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProSerIleLeuSerMet	370
QY	1300	ATTTCATGTCGCGCAAGCACACTCCTCTACACAGCTGTATTGTTTGCACCCCTTTGACAAATG	1359
Db	371	IleHisProGlnLeuLeuThrProValProSerLeuValPheThrCysValMetThrLeu	390
QY	1360	ATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCTCAGTTTGGCAGGTGG	1419
Db	391	MetTyrAlaPheSerArgAspIlePheSerIleIleAsnPhePheSerPhePheAsnTrp	410
QY	1420	CTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATG	1479
Db	411	LeuCysValAlaLeuAlaIleIleGlyMetMetTrpLeuArgPheLysLysProGluLeu	430
QY	1480	CATCGTCCCTTTCAAGGTGCCACTGTTTCATCCACGCTTTGTTTTCTCTCATGCTCTTCT	1539
Db	431	GluArgProIleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPhe	450
QY	1540	ATGGTTGCCCTTTCCCTCTATTTCGGGACCCCATTTAGTACAGGATTGGCTTCGTCATCACT	1599
Db	451	LeuIleAlaValSerPheTrpLysThrProLeuGluCysGlyIleGlyPheAlaIleIle	470
QY	1600	CTGACTGGAGTCCCTCGGTATTATCTCTTTATTATATGGGACAGAAACCCAGGTGGTTT	1659
Db	471	LeuSerGlyLeuProValTyrPhePheGlyValTrpTrpLysAsnLysProLysTrpIle	490
QY	1660	AGAATAATGTCAGAGAAAATAACCGAACATTAACAATAATACTGGAAGTTGTACCAGAA	1719
Db	491	LeuGlnValIlePheSerValThrValLeuCysGlnLysLeuMetGlnValValProGln	510
QY	1720	GAA 1722	
Db	511	Glu 511	

RESULT 6
ABB12114
ID ABB12114 standard; peptide; 507 AA.
XX
AC
XX
AC
XX
DT 11-JAN-2002 (first entry)
XX
DE Human 4F2 light chain homologue, SEQ ID NO:2484.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PN
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR N-PSDB; ABA09358.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 20; page 308-309; 1963pp; English.

Sequences ABB10981-ABBI12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides, or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth factor activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and

CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

xx Sequence 507 AA;

Alignment Scores: 1.14e-119 Length: 507
Pred. No.: 1140.00 Matches: 218
Percent Similarity: 66.17% Conservative: 95
Best Local Similarity: 46.09% Mismatches: 156
Query Match: 27.67% Indels: 4
DB: 22 Gaps: 3

US-09-667-170A-440 (1-2239) x ABB12114 (1-507)

QY	247	CCTTCCCTGGCAACAGGAGCCACCTGGGCGAGGAGCGCTTTTCAGGAAGAGACGCGCTTT	306
Db	14	ProValAlaGluGluLysGluGluAlaArgGluLysIleMetAlaAlaLysArgAlaAsp	33
QY	307	TCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTA	366
Db	34	GlyAlaAlaProAlaGlyGluGly---GluGlyValThrLeuGlnGlyAsnIleThrLeu	52
QY	367	CTGAGGGAGTCTCCATTATCATTTGGCACCACATATTGGAGCAGGAATCTTCATCTCTCCT	426
Db	53	LeuLysGlyValAlaValIleValAlaIleMetGlySerGlyIlePheValThrPro	72
QY	427	AAGGGCGTGTCCAGAACACAGCGTGGGCGATGTCTCTGACCATCTGGACGGTGTGT	486
Db	73	ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValThrAlaAlaCys	92
QY	487	GGGGTCTGTCACTATTTGGAGCTTTGTCTTATGTGTAATGGGAACAACATAAAGAAA	546
Db	93	GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleSerLys	112
QY	547	TCTGGAGGTCAATACACATATATTTTGAAGTCTTTGGTCCATTACCAAGCTTTTGTACGA	606
Db	113	SerGlyGlyAspTyrAlaTyrMetLeuAspValTyrGlySerProAlaPheLeuLys	132
QY	607	GTCTGGGTGGAACCTCCTCATAATACGCCCTGCAGCTACTGTCTGTGATATCCCTGGCATTT	666
Db	133	LeuTrpIleGluLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe	152
QY	667	GGACGCTACATTTCTGGAACCATTTTATTTCAATGTGAATCCCTGAACCTTGGCATCAAG	726
Db	153	AlaThrTyrLeuLeuLysProLeuPheProThrCysProValProGluGluAlaAlaLys	172
QY	727	CTCATTTACAGCTGGGCATAACTGTAGTGTGCTTAAATAGCATGAGTGTGAGCTGG	786
Db	173	LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla	192
QY	787	AGCGCCCGGATCCAGATTTCTTAACCTTTTGCAGGCTCACAGCAATCTGTATAATTATA	846
Db	193	AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuIleIle	212
QY	847	GTCCCTGGAGTTATGCAGCTAATTAAGGTCAACAGCAGAACTTTAAAGACGCTTTTCA	906
Db	213	LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer	232
QY	907	-----GGAAGAGATTCAAGTATTACGCGGTGGCCACTGGCTTTTATTTATGGAATGTAT	960
Db	233	PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe	252

QY	961	GCATATGCTGGCTGGTGTCTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCCCTGAAAA	1020
Db	253	AlaTyrGlyGlyTyrAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg	272
QY	1021	ACCATTCCTCCTTGCATATATGATATCCATGGCCATTGTACCATTTGGCTATGTGCTGACA	1080
Db	273	AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr	292
QY	1081	AATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGA	1140
Db	293	AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla	312
QY	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCGATCTTTGTGGC	1200
Db	313	ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly	332
QY	1201	CTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTTGTCTCTCCAGGTTATCTATGTT	1260
Db	333	LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPhePheVal	352
QY	1261	CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCCGACGACACT	1320
Db	353	GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr	372
QY	1321	CCTCTACAGCTGTATTATTGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGAC	1380
Db	373	ProValProSerLeuValPheThrCysValMetThrLeuPheThrAlaPheSerLysAsp	392
QY	1381	CTCGACAGCTCTTTGAATTTCTCAGTTTTCAGGTTGCTGCTGCTGCTGCTGCTGCTGCT	1440
Db	393	IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle	412
QY	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATGCTGCTGCTGCTGCTGCT	1500
Db	413	IleGlyMetIleThrLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn	432
QY	1501	CTGTTTCATCCCGCTTTGTTTTCCTTTCACATGCTCTTTCATGGTGGCTGCTGCTGCT	1560
Db	433	LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp	452
QY	1561	---TCGGACCCCATTTAGTACAGGGATTGGCTTCGTCATCATCTGCTGAGTCCCTGCG	1617
Db	453	LysThrThrProTrpSerValAlaSerAspPheThrIleIleLeuSerGlyLeuProVal	472
QY	1618	TATTATCTCTTTTATATATATGGGACAAAGAAACCCAGGTGG	1656
Db	473	TyrPhePheGlyValTrpTrpLysAsnLysProLysTrp	485

RESULT 7

AA41159	AA41159 standard; Protein; 511 AA.
XX	AA41159;
AC	AA41159;
XX	24-JAN-2000 (first entry)
DT	Human amino acid permease homolog (AAPH).
XX	Amino acid permease homolog; AAPH; cancer; inflammatory disorder; AIDS;
KW	autoimmune disorder; anemia; asthma; atherosclerosis; diabetes mellitus;
KW	multiple sclerosis; osteoarthritis; osteoporosis; rheumatoid arthritis;
KW	ulcerative colitis; infection; cell proliferation disorder; human.
XX	Homo sapiens.
OS	US5981242-A.
PN	09-NOV-1999.
XX	03-FEB-1999; 99US-0243920.
PF	06-MAY-1998; 98US-0073362.
XX	
PR	
XX	

PA (INCY-) INCYTE PHARM INC.
XX Yue H, Corley NC, Hillman JL;
XX WPI; 1999-633325/54.
DR N-PSDB; AAZ23240.
XX
PT New human amino acid permease homolog, useful in the diagnosis,
PT treatment and prevention of cancer, inflammatory/autoimmune disorders
PT and cell proliferation disorders -
XX
PS Claim 1; Fig 1A-F; 32pp; English.
XX
CC This represents a human amino acid permease homolog (AAPH). The AAPH
CC nucleic acid, protein, AAPH (ant)agonists are useful in the diagnosis,
CC treatment and prevention of cancer such as adenocarcinoma, leukemia,
CC lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the
CC adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall
CC bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung,
CC muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands,
CC skin, spleen, testis, thymus, thyroid or uterus; inflammatory/autoimmune
CC disorders such as AIDS, Addison's disease, adult respiratory distress
CC syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia,
CC asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune
CC thyroiditis, bronchitis, cholecystitis, contact dermatitis, diabetes
CC mellitus, Crohn's disease, atopic dermatitis, dermatomyositis,
CC emphysema, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid
CC arthritis, ulcerative colitis, viral, bacterial, fungal, parasitic,
CC protozoal infections; and cell proliferation disorders such as actinic,
CC keratosis, cirrhosis, arteriosclerosis, myelofibrosis, and primary
CC thrombocythemia.
XX

SQ Sequence 511 AA;

Alignment Scores:

Pred. No.:	1.05e-118	Length:	511
Score:	1131.50	Matches:	214
Percent Similarity:	68.00%	Conservative:	109
Best Local Similarity:	45.05%	Mismatches:	149
Query Match:	27.46%	Indels:	3
DB:	20	Gaps:	2

US-09-667-170A-440 (1-2239) x AAY41159 (1-511)

QY	310	GGAAGAGACGGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTG	369
Db	21	GlyAspGlyAlaSerProGlyProGluGlnValLysLeuLysLysGluLeuSerLeu	40
QY	370	AGGGAGTCTCCATATATCATTTGGCACCACATCATTTGGAGCAGGAATCTTCTCTCCTAAG	429
Db	41	AsnGlyValCysLeuIleValGlyAsnMetIleGlySerGlyIlePheValSerProLys	60
QY	430	GGCGTGTCCAGAACACGGGCGGCGGCGATGTCTCTGACCATCTGGACGGTGTGGG	489
Db	61	GlyValLeuIleTyrSerAlaSerPheGlyLeuSerLeuValIleTrpAlaValGlyGly	80
QY	490	GTCCTGTCACTATTGGAGCTTTTGTCTTATGCTGAATTGGGAACAACATAAAGAAATCT	549
Db	81	LeuPheSerValPheGlyAlaLeuCysTyrAlaGluLeuGlyThrIleLysLysSer	100
QY	550	GGAGGTCAATACACATATATTTTGAAGTCTTTTGGTCCATTACCAGCTTTTGTACGAGTC	609
Db	101	GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu	120
QY	610	TGGGTGGAACCTCCTCATATACGCCCTGCAGCTACTGCTGFGATATCCCTGGCATTTGGA	669
Db	121	TrpThrSerLeuLeuIleGluProThrSerGlnAlaIleIleAlaIleThrPheAla	140
QY	670	CGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACCTTGGCATCAAGCTC	729
Db	141	AsnTyrMetValGlnProLeuPheProSerCysPheAlaProTyrAlaAlaSerArgLeu	160
QY	730	ATTACAGCTGTGGGCATAACTGTAGTGGTCTCTAAATACCATGAGTGTCAAGTGGAGC	789

Db	161	LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly	180
QY	790	GCCGGATCCAGATTTTCTTAACCTTTTGCAGCTCACAGCAATTTCTGATAATATAGTC	849
Db	181	ThrLeuValGlnAspIlePheTrpTyrAlaLysValLeuAlaLeuIleAlaValIleVal	200
QY	850	CCTGGAGTTATGCAGCTAATTAAGGTCAAACGCCAGAACTTTAAAGACGCCCTTTTCAGGA	909
Db	201	AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly	220
QY	910	AGAGATTCAAGTATTACGGCGTTGCCACTGGCTTTTATTATGGAATGTATGCATATGCT	969
Db	221	SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer	240
QY	970	GGCTGGTTTTTACCTCAACTTTGTACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTC	1029
Db	241	GlyTrpAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro	260
QY	1030	CTTGCATATATATATCCATGGCCCATTTGTACCATTTGGCTATGTGTGACAAATGTGGCC	1089
Db	261	LeuSerIleGlyIleSerMetProIleValThrIleIleTyrIleLeuThrAsnValAla	280
QY	1090	TACTTTACGACCATTAAATGCTGAGGAGCTGTCTTCAAAATGCAGTGGCAGTGCCTTT	1149
Db	281	TyrTyrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe	300
QY	1150	TCTGAGCGGTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTGGCCCTCTCCTGC	1209
Db	301	AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys	320
QY	1210	TTTGGCTCCATGAACGGTGTGTGTGTGTCTCTCCAGGTTATTTCTATGTTGGTCTCGA	1269
Db	321	PheGlyGlyLeuAsnAlaSerIleValAlaAlaAlaSerArgLeuPheValGlySerArg	340
QY	1270	GAGGTCACCTTCCAGAAATCTCTCCATGATTCATGTCGCGCAAGCACACTCTCTACCA	1329
Db	341	GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro	360
QY	1330	GCTGTTATTGTTTGGACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACCTCGACAGT	1389
Db	361	SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln	380
QY	1390	CTTTTGAATTTTCTCAGTTTTCAGGTTGGCTTTTATTGGCTGGCAGTTGCTGGGCTG	1449
Db	381	LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln	400
QY	1450	ATTTATCTTCGATACAAATGCCAGATATGCATCGCTCTTCAAGGTGCCACTGTTTCATC	1509
Db	401	LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe	420
QY	1510	CCAGCTTTGTTTCTTCACATGCCCTCTTCATGGTGGCTTCCCTCTTATTCGGACCCA	1569
Db	421	ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr	440
QY	1570	TTTAGTACAGGGATTGGCTTCGTCATCATCTGACTGGAGTCCCTGGCTATTATCTCTTT	1629
Db	441	IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle	460
QY	1630	ATT-----ATATGGACAAAGAAACCCAGGTGGTTTGAATAATGTACAGAAAAATAACC	1683
Db	461	IleArgValProGluHisLysArgProLeuTyrLeuArgIleValGlySerAlaThr	480
QY	1684	AGAACATTACAAATAATA---CTGGAAGTTGTACAGAAAGAAGAT	1725
Db	481	ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp	495

RESULT 8

AAY27076

ID AAY27076 standard; Protein; 511 AA.

XX

AC AAY27076;

XX

[illegible]


```
QY 1570 TTTAGTACAGGATTGGCTTCGTCACTCACTGACTGGAGTCCCTGGCTATTATCTCTTT 1629
Db 441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle 460
QY 1630 ATT-----ATATGGGACAAAGAAACCCAGGTGTTTGAATAATATGTCAGAGAAATAACC 1683
Db 461 IleArgValProGluHisLysArgProLeuTyrLeuArgArgIleValGlySerAlaThr 480
QY 1684 AGAACATTACAAATAATA---CTGGAAGTTGTACCAAGAAAGAT 1725
Db 481 ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp 495
RESULT 10
ID ABB58292 standard; Protein; 505 AA.
XX
AC ABB58292;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 1668.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL02395.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 1668; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 505 AA;
Alignment Scores:
Pred. No.: 1.16e-116 Length: 505
Score: 1113.50 Matches: 208
Percent Similarity: 64.50% Conservative: 110
Best Local Similarity: 42.19% Mismatch: 170
Query Match: 27.03% Indels: 5
DB: 22 Gaps: 2
US-09-667-170A-440 (1-2239) x ABB58292 (1-505)
```

```
QY 235 AACGGGAGGCTGCCCTTCCTGGGCAACAAGGACACCTGGGAGGACGCTTTTCAGGA 294
Db 11 SerSerArgLysSerSerIleValAsnGlyAsnGlyAspAlaSerAlaLysLeuThrAsn 30
QY 295 AGAGACGCCTTTTCAGGAAGAGACAGCCCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGG 354
Db 31 GlyAspGlyAspGlyGlyAsp-----GlyGlyGlyGluValThrLeuLysAla 47
QY 355 AAAGTCACCTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATC 414
Db 48 LysMetSerLeuLeuAsnGlyCysThrValIleValGlySerIleIleGlySerGlyIle 67
QY 415 TTCATCTCTCTAAGGGCGTGTCTCCAGAACACAGCGGTGGGCATGTCTCTGACCATC 474
Db 68 PheValSerProThrGlyValLeuMetTyrThrGlySerValAsnLeuAlaLeuIleVal 87
QY 475 TGGACGGTGTGGGGTCTCTCACTATTTTGGAGCTTTTGTCTTATGTCTGAATTTGGGAACA 534
Db 88 TrpValIleSerGlyLeuPheSerMetValGlyAlaTyrCysTyrAlaGluLeuGlyThr 107
QY 535 ACTATAAGAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACCA 594
Db 108 MetIleThrLysSerGlyAlaAspTyrAlaTyrIleMetGluThrPheGlyProPheMet 127
QY 595 GCTTTTGTACGAGTCTGGTGGAACTCTCATATAATACGCCCTGCAGCTACTGCTGTGATA 654
Db 128 AlaPheIleArgLeuTrpIleGluCysMetIleValArgProCysSerSerGlnAlaIleVal 147
QY 655 TCCCTGGCATTTGGACGTACATCTCTGGAACCATATTTTATCAATGTGAAATCCCTGAA 714
Db 148 AlaLeuThrPheSerThrTyrValLeuLysProphePheProGluCysThrProProGlu 167
QY 715 CTTCGGATCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGAGTCTCTAAATAGCATG 774
Db 168 AspSerAlaArgLeuLeuAlaValCysCysIleLeuValLeuThrLeuIleAsnCysTrp 187
QY 775 AGTGTGAGCTGGAGCGCCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATT 834
Db 188 AspValLysTrpAlaThrAlaValGlnAspIlePheThrTyrAlaLysLeuLeuAlaLeu 207
QY 835 CTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAACGACGAGAACTTTAAA 894
Db 208 PheIleIleAlaThrGlyValTyrGlnLeuTyrLeuGlyAsnThrGlnTyrPhe--- 226
QY 895 GACGCCTTTTCAGGAAGAGATTCAAGTATTACCGGTTGCCACTGGCTTTTATTATGGA 954
Db 227 ---ThrPheGluAsnThrAspThrLysValThrSerIleAlaLeuSerPheTyrSerGly 245
QY 955 ATGTATGCATATGTGGCTGGTGTTCCTACCTCAACTTTGTACTGAAGAGTAGAAAACCCCT 1014
Db 246 LeuPheAlaTyrAsnGlyTrpAsnTyrLeuAsnPheIleIleGluGluLeuLysAspPro 265
QY 1015 GAAAAAACCATTTCCCTTGAATATGTATATCCATGGCCATTTGTACCATTTGGTATGTG 1074
Db 266 ValLysAsnLeuProArgAlaIleAlaIleSerCysThrLeuValThrIleValTyrVal 285
QY 1075 CTGACAAATGTGGCTACTTTACGACCATTAATGTGTGAGGAGCTGTCTTCAAAATGCA 1134
Db 286 MetAlaAsnValSerPheTyrThrIleLeuSerProAspGluValMetGlySerSerAla 305
QY 1135 GTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTCTCCGATCTTT 1194
Db 306 ValAlaValThrTyrAlaGluArgAlaPheGlyMetLeuAlaTrpThrIleProValPhe 325
QY 1195 GTTGGCTCTCTCTGCTTTGGCTCCCATGAACGGTGGTGTGTGTGTGTGTGTGTGTGTGT 1254
Db 326 ValAlaLeuSerThrPheGlyAlaValAsnGlyIleLeuLeuThrSerSerArgLeuPhe 345
QY 1255 TATGTTGCGTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTTCATGTCGCAAG 1314
Db 346 TyrAlaGlyAlaAsnAsnGlyGlnMetProGluIleLeuThrMetIleGlnIleGlnArg 365
QY 1315 CACACTCTCTTACCAGCTGTATTATGTTTGTGACCCCTTTGACAAATGATAATGCTCTCTCT 1374
```


Db 366 PheThrProThrProAlaValLeuAlaMetAlaLeuLeuSerMetLeuThrVal 385
QY 1375 GGACACCTCGACAGCTTTTGAATTTCTCAGTTTGGCCAGGTGGCTTTTATTTGGCTG 1434
Db 386 SerAspIlePheAlaLeuIleAsnTyrValGlyPheAlaThrTrpLeuSerIleGlyVal 405
QY 1435 GCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAG 1494
Db 406 AlaValLeuCysLeuProThrLeuArgTrpAlaGlnProAsnLeuProArgProIleArg 425
QY 1495 GTGCCACTGTTTCATCCAGCTTTGTTTCTTCACATGCCCTCTTCATGGTTGCCCTTTCC 1554
Db 426 ValProMetValPheProIleValTyrLeuIleAlaThrIlePheValThrValValPro 445
QY 1555 CTCTATTTCGGACCCATTTAGTACAGGATTTGGCTTCGTCACTCACTCTGAGTGGAGTCCCT 1614
Db 446 MetTyrAlaSerProValGluThrGlyTyrGlyIleLeuMetIleLeuSerSerIlePro 465
QY 1615 GCGTATTATCTCTTTATTTATATGGGACAGAAACCCAGGTGGTTTGAATAATGTCAGAG 1674
Db 466 ValTyrLeuValPheIleAlaTrpLysAsnLysProIleTrpPheGlnLysThrMetGly 485
QY 1675 AAAATAACCCAGAACATTTACAAATAATATCTGGAAGTTGTA 1713
Db 486 GlyLeuThrGlnValLeuGlnLysLeuMetMetValVal 498

RESULT 11
ID AAB72393 standard; Protein; 533 AA.
XX AAB72393;
AC AAB72393;
XX AAB72393;
DT 24-MAY-2001 (first entry)
XX Rat L-type amino acid transporter 2 (LAT2) amino acid sequence.
DE Neutral amino acid transporter; amino acid transport activating factor;
KW 4F2hc; 4F2 heavy chain; drug permeation; poison; rat; LAT2;
KW L-type amino acid transporter 2.
XX
OS Rattus sp.
XX
XX JP2000342270-A.
XX
XX 12-DEC-2000.
XX
XX 04-JUN-1999; 99JP-0158571.
XX
XX 04-JUN-1999; 99JP-0158571.
XX
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI; 2001-260288/27.
DR N-PSDB; AAF77276.
XX
PT A neutral amino acid transporter having broad substrate selectivity
PT useful for in vitro analysis of permeation of drugs and poisons through
PT cell membranes, blood and tissues.
XX
PS Claim 4; Page 17-18; 37pp; Japanese.
XX
CC This invention relates to a neutral amino acid transporter having broad
CC substrate selectivity to neutral amino acids, and which is inactivated by
CC an amino acid transport activating factor 4F2hc (4F2 heavy chain). The
CC neutral amino acid transporter and its gene are useful in vitro
CC analysis of permeation of drugs and poisons through cell membrane, blood
CC and tissues. The present sequence represents the rat L-type amino acid
CC transporter 2 (LAT2).
XX

SQ Sequence 533 AA;

Alignment Scores:

Pred. No.: 5.03e-116 Length: 533
Score: 1108.00 Matches: 224
Percent Similarity: 63.04% Conservative: 100
Best Local Similarity: 43.58% Mismatches: 168
Query Match: 26.89% Indels: 22
DB: 22 Gaps: 5

US-09-667-170A-440 (1-2239) x AAB72393 (1-533)

QY 202 ATCTCCAAAGGAGGTTACCTGCAGGGAATGTTAACGGGAGGCTGCCTTCCCTGGGCAAC 261
Db 1 MetGluLysGlyThrArgGlnArgAsnAsnThrAlaLysAsnHisProAspArgGlySer 20
QY 262 AAGGAGCCACCTGGGCGAGGAGCGCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGACGCC 321
Db 21 AspThrSerProGluAlaGluAlaSerSerGlyGlyGly----- 34
QY 322 TTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCC 381
Db 35 -----ValAlaLeuLysLysGluIleGlyLeuValSerAlaCysGly 48
QY 382 ATTATCATTTGGCACCACCATTTGAGCAGGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 441
Db 49 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 68
QY 442 AACACGGGCGAGCGTGGGCGATCTCTCTGACCATCTGGACGCTGTGGGGTCTGTCTACTA 501
Db 69 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyValIleThrAla 88
QY 502 TTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTAC 561
Db 89 ValGlyAlaLeuCysTyrAlaGluLeuGlyValTrpIleProLysSerGlyGlyAspTyr 108
QY 562 ACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTGTACGAGTCTGGTGGAACTC 621
Db 109 SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 128
QY 622 CTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTG 681
Db 129 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 148
QY 682 GAACCATTTTATTCAATGTGAATCCCTGAACTTCCGATCAAGCTCATACAGCTGTG 741
Db 149 GlnProLeuPheProThrCysPheProProGluSerGlyLeuArgLeuLeuAlaIle 168
QY 742 GGCATAACTGTAGTGATGCTCTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAG 801
Db 169 CysLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 188
QY 802 ATTTTCTTAACCTTTTCAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATG 861
Db 189 AspIlePheThrAlaGlyLysLeuLeuAlaLeuAlaLeuIleIleMetGlyValVal 208
QY 862 CAGCTAATTAAGGTCACACGAGCAACTTT-----AAAGACGCCCTTTTCA 906
Db 209 GlnIleCysLysGlyGlu-----PhePheTrpLeuGluProLysAsnAlaPheGlu 225
QY 907 GGA---AGAGATTCAAGTATTACGGGGTTGCCACTGGCTTTTATTTATGGAATGTATGCA 963
Db 226 AsnPheGlnGluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAla 245
QY 964 TATGCTGGCTGGTTTACCTCAACTTTTGTACTGAAGAGTAGAAACCCCTGAAAAAAC 1023
Db 246 TyrGlyGlyTrpAsnPheLeuAsnTyrValThrGluGluLeuValAspProTyrLysAsn 265
QY 1024 ATTCCTTGGCAATATGATATCCATGGCATTTGTCACCATTTGGCTATGTGCTGACAAAT 1083
Db 266 LeuProArgAlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsn 285
QY 1084 GTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTG 1143
Db 286 IleAlaTyrValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaVal 305

QY 1144 ACCTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCAGATCTTTGTTGCCCTC 1203
Db 306 ThrPheGlyGluLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeu 325
QY 1204 TCCTGCTTTGGCTCCATGAACGGTGGTGTGCTGCTCCAGGTATTCTATGTTGCG 1263
Db 326 SerThrPheGlyGlyValAsnGlySerLeuPheThrSerArgLeuPhePheAlaGly 345
QY 1264 TCTCGAGAGGGTCACTCCAGAAATCCTCCATGATTCATGTCGCGCAAGCACACTCCT 1323
Db 346 AlaArgGluGlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrPro 365
QY 1324 CTACAGAGTGTATTGTTTGCACCTTTGACAATGATAATGCTCTTCTGAGACCTC 1383
Db 366 IleProAlaLeuLeuPheThrCysLeuSerThrLeuLeuMetLeuValThrSerAspMet 385
QY 1384 GACAGTCTTTTGAATTTCTCAGTCTTTGCCAGGTGGCTTTTATTGGGCTGGCAGTTGCT 1443
Db 386 TyrThrLeuIleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAla 405
QY 1444 GGGCTGATTATCTTCGATACAAATGCCAGATATGTCATCGTCTTTCAGGTGCCACTG 1503
Db 406 GlyGlnIleValLeuArgTrpLysLysProAspIleProArgProIleLysIleSerLeu 425
QY 1504 TTATCCAGCTTTGTTTCTTCATGCTCTTCATGCTTTCATGCTTTCCTCTTATTCG 1563
Db 426 LeuPheProIleIleTyrLeuLeuPheTrpAlaPheLeuIlePheSerLeuTrpSer 445
QY 1564 GACCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTAT 1623
Db 446 GluProValValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPhe 465
QY 1624 CTCTTTATATATATGGACAAAGAACCCAGGTGGTGGTTTACAATAATGTGACAGAAATAACC 1683
Db 466 LeuGlyValTyrTrpGlnHisLysProLysCysPheAsnAspPheIleGluSerLeuThr 485
QY 1684 AGAACATTAACAATAATACTGGAAGTTGTA---CCAGAGAA 1722
Db 486 LeuValSerGlnLysMetCysValValValTyrProGlnGlu 499

RESULT 12
ABB71209
ID ABB71209 standard; Protein; 499 AA.
XX ABB71209;
AC ABB71209;
XX ABB71209;
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 40419.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX N-PSDB; ABL15312.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -
XX Disclosure; SEQ ID NO 40419; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 499 AA;
Alignment Scores:
Pred. NO.: 8.16e-116 Length: 499
Score: 1106.00 Matches: 218
Percent Similarity: 65.57% Conservative: 102
Best Local Similarity: 44.67% Mismatches: 156
Query Match: 26.84% Indels: 12
DB: 22 Gaps: 3
US-09-667-170A-440 (1-2239) x ABB71209 (1-499)
QY 265 GAGCCACCTGGGAGGAGCGCTTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGCGCTTT 324
Db 22 GluProAsnAsnSerThrAlaAspSerGlySerGlnGly----- 34
QY 325 TCAGGAAGAGAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTACTGAGGGAGTCTCCATT 384
Db 35 SerGly-----ValLysLeuLysLysGlnIleGlyLeuLeuAspGlyValAlaIle 51
QY 385 ATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAAC 444
Db 52 IleValGlyValIleValGlySerGlyIlePheValSerProLysGlyValLeuLysPhe 71
QY 445 ACGGGCAGCGTGGGCGATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGCTACATATT 504
Db 72 SerGlySerIleGlyGlnSerLeuIleValTrpValLeuSerGlyValLeuSerMetVal 91
QY 505 GGAGCTTTGCTTATGCTGAATTTGGGAACAACATAAAGAAATTCGAGGTCAATACACA 564
Db 92 GlyAlaLeuCysTyrAlaGluLeuGlyThrMetIleProLysSerGlyGlyAspTyrAla 111
QY 565 TATATTTTGGAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGTGGAACTCCTC 624
Db 112 TyrIleGlyThrAlaPheGlyProLeuProAlaPheLeuTyrLeuTrpValAlaLeuLeu 131
QY 625 ATAATACGCCCTGACGCTACTGCTGTATATCCCTGGCATTTGGACGCTACATTTCTGGAA 684
Db 132 IleLeuValProThrGlyAsnAlaIleThrAlaLeuThrPheAlaIleTyrLeuLys 151
QY 685 CCATTTTATTCAATGTGAATCCCTGAACTTCGATTCGATTCAGCTCAAGCTACAGCTGGGC 744
Db 152 PropheTrpProSerCysAspAlaProIleGluAlaValGlnLeuLeuAlaAlaMet 171
QY 745 ATAAGTGTAGTGGTCTCTAAATAGCATGAGTGTGAGTGGAGCGCGGATCCAGATT 804
Db 172 IleCysValLeuThrLeuIleAsnCysTyrAsnValLysTrpValThrArgValThrAsp 191
QY 805 TTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAG 864
Db 192 IlePheThrGlyThrLysValValAlaLeuLeuValIleValGlyAlaGlyValTrpTrp 211
QY 865 CTAATTAAGGTCAACACGCAAGAACTTTTAAAGACGCGCTTTTTCAGGAAGAGATTCAAGTATT 924
Db 212 LeuPheAspGlyAsnThrGluHisTrpAspAsnPropheSerGlyGlyLeuGlnAspPro 231
QY 925 ACGGGTGGCCACTGCTTTTATTATTGGAATGATGATATGCTGGCTGTTTACCTC 984

Db 232 GlyTyrIleAlaLeuAlaPheTyrSerGlyLeuPheSerTyrSerGlyTrpAsnTyrLeu 251
QY 985 AACTTTGTTACTGAAGAAGTAGAAAACCTGAAAAAACCAATCCCTTGCAATATGTATA 1044
Db 252 AsnPheValThrGluGluLeuLysAspProTyrArgAsnLeuProLysAlaIleCysIle 271
QY 1045 TCCATGGCCATTGTCACCAATTGGCTATGTGCTGACAAAATGTGGCCTACTTTACGACCAATT 1104
Db 272 SerMetProValThrValIleTyrMetIleThrAsnIleAlaTyrPheSerValLeu 291
QY 1105 AATGCTGAGGAGTGTCTGCTTTCAAATGAGTGGCAGTGACCTTTTCTGAGCGGCTACTG 1164
Db 292 SerProAspGluIleLeuSerSerAspAlaValAlaValThrPheGlyAspLysMetLeu 311
QY 1165 GGAAATTTCTCAATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTGGCTCCATGAAC 1224
Db 312 GlyTyrMetSerTrpIleMetProPheAlaValAlaCysSerThrPheGlySerLeuAsn 331
QY 1225 GGTGGTGTGTGCTCTCCAGGTTAFTCTATGTTGGTCTCGAGAGGGTCACCTTCCA 1284
Db 332 GlyAlaIlePheAlaSerSerArgLeuPhePheValGlyAlaArgAsnGlyHisLeuPro 351
QY 1285 GAAATCCTCTCCATGATTTCATGTCCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTG 1344
Db 352 AlaAlaIleSerLeuIleAsnValAsnCysLeuThrProValProSerLeuIlePheLeu 371
QY 1345 CACCCTTTGACATGATAATGCTCTTCTCTGAGACCTCGACAGTCTTTTGAATTCCTC 1404
Db 372 GlyValLeuThrLeuLeuLeuPheIleGluAspValTyrValLeuIleAsnTyrVal 391
QY 1405 AGTTTGGCAGGTGGCTTTTATTGGGTGGCAGTTGTGGGTGATTTATCTTCGATAC 1464
Db 392 SerTyrValGluAlaLeuPheThrLeuIleSerValSerGlyLeuLeuTrpMetArgTyr 411
QY 1465 AAATGCCAGATATGATCGTCTTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCC 1524
Db 412 LysGlnProLysThrGluArgProIleLysValAsnLeuAlaLeuProIleIleTyrLeu 431
QY 1525 TTCACATGCTCTTTCATGGTTGCCCTTTCCCTCTATTCGGACCCATTAGTACAGGGATT 1584
Db 432 IleValCysLeuPheLeuValIlePheSerCysThrGlnThrProTyrValValGlyIle 451
QY 1585 GGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTTATTATATGGGACAAG 1644
Db 452 GlyThrIleIleLeuSerGlyIleProValTyrTyrLeuThrIle-----HisLys 469
QY 1645 AAACCCAGGTGGTTTAGAATAATGTACAGAGAAATAACACAGACATTACAAATAATACTG 1704
Db 470 ProValLysTrpLeuAlaAspThrSerGlnAlaIleAsnLeuTrpCysSerLysPhePhe 489
QY 1705 GAAGTTGTACCAAGAAGATAAG 1728
Db 490 IleCysMetProAsnGlnGluLys 497

RESULT 13

AAB43136
ID AAB43136 standard; Protein: 535 AA.

XX AAB43136;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2900 polypeptide sequence SEQ ID NO:5800.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC77345.

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -

PS Claim 11; Page 4966-4968; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX

CC sequences have activities such as: cytostatic; hepatotropic; vulnery;

CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;

CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;

CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;

CC antidiabetic; hypotensive; dermatological; immunosuppressive;

CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;

CC antithyroid; and antianaemic. The sequences can be used for determining

CC the presence of or predisposition to, or preventing or treating

CC pathological conditions associated with an ORFX-associated disorder. The

CC nucleic acids can be used to express ORFX proteins in gene therapy

CC vectors. The proteins and nucleic acids may be used to treat cancers,

CC proliferative disorders, neurodegenerative disorders, osteoarthritis,

CC graft vs host disease, cardiovascular disease, diabetes mellitus,

CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,

CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,

CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,

CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance

CC coagulation; to inhibit thrombosis; and as a contraceptive.

SQ Sequence 535 AA;

Alignment Scores:

Pred. No.: 2.76e-115 Length: 535
Score: 1101.50 Matches: 220
Percent Similarity: 65.52% Conservative: 103
Best Local Similarity: 44.62% Mismatches: 161
Query Match: 26.74% Indels: 9
DB: 21 Gaps: 4

US-09-667-170A-440 (1-2239) x AAB43136 (1-535)

QY 262 AAGGAGCCACCTGGCAGGACGCCTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGACGCC 321

Db 13 LysLysHisProGly-----GlyGlyGluSerAspAlaSerProGluAla 27

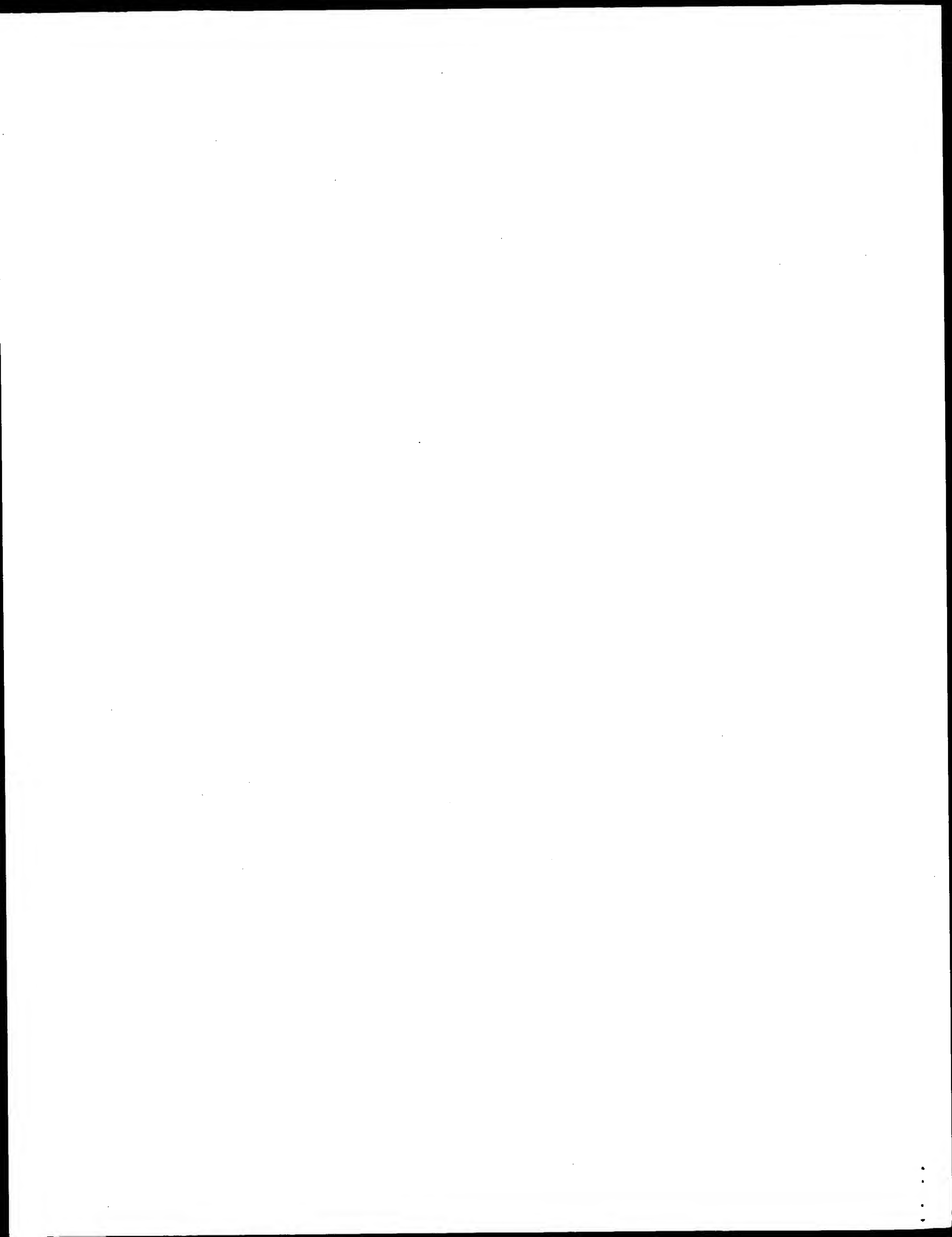
QY 322 TTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTTACTGAGGGGAGTCTCC 381

Db 28 GlySerGlyGlyGlyValAlaLeuLysLysGluLeuGlyLeuValSerAlaCysGly 47
QY 382 ATTATCATTTGGCACCATTATGGAGCAGGAATCTCATCTCTCCTAAGGGGTGCTCCAG 441
Db 48 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 67
QY 442 AACACGGGACGCTGGGATGCTCTGACCATCTGACGGGTGTGTGGGTCCTCACTA 501
Db 68 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyPheIleThrVal 87
QY 502 TTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCAATAC 561
Db 88 ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 107
QY 562 ACATATATTTTGAAGCTCTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGTGAACATC 621
Db 108 SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 127
QY 622 CTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGCTACATCTGTG 681
Db 128 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 147
QY 682 GAACCATTTTATTCAATGTGAATCCCTGAACCTGCGATCAAGCTCATACAGCTGTG 741
Db 148 GlnProLeuPheProThrCysPheProProGluSerGlyLeuArgLeuLeuAlaIle 167
QY 742 GGCATAAAGTGTAGTGTGCTGCTAAATAGCATGAGTGTGAGTGGAGCCGGATCCAG 801
Db 168 CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 187
QY 802 ATTTCTTAACTTTTGCAGACTCACAGCAATCTGATAATTATAGTCCCTGGAGTTATG 861
Db 188 AspIlePheThrAlaGlyLysLeuLeuAlaLeuAlaLeuIleIleMetGlyIleVal 207
QY 862 CAGCTAATTAAGGTCAA-----ACGCAGAACTTTAAAGACGCTTTTTCAGGA--AGA 912
Db 208 GlnIleCysLysGlyGlyLeuTrpPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln 227
QY 913 GATTCAAGTATTACGGGGTGGCCACTGGCTTTTATTATGGAATGTATGATATGCTGGC 972
Db 228 GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaTyrGlyGly 247
QY 973 TGGTTTACCTCAACTTTGTACTGAAGAAGTAGAAACCCCTGAAACCAACCATTCCTCCTT 1032
Db 248 TrpAsnPheLeuAsnTyrValThrGluLeuValAspProTyrLysAsnLeuProArg 267
QY 1033 GCAATATGTATATCCATGCCCATTTGTACCATTTGGCTATGTCTGACAAATGTGGCCTAC 1092
Db 268 AlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsnValAlaTyr 287
QY 1093 TTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCAGTACCTTTTCT 1152
Db 288 ValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaValThrPheGly 307
QY 1153 GAGCGGCTACTGGGAAATTTCTCATAGCAGTTCGGATCTTTGTTGCCCTCTCCTGCTTT 1212
Db 308 GluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeuSerThrPhe 327
QY 1213 GGCTCCATGAACGGTGGTGTGTTTGTGCTCTCCAGGTATTCTATGTTGGCTCTCGAGAG 1272
Db 328 GlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGlyAlaArgGlu 347
QY 1273 GGTCCACTTCCAGAAATCTCTCCATGATTCATGTCGGAACACACACTCTCTACCACT 1332
Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrProIleProAla 367
QY 1333 GTTATTGTTTTCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTT 1392
Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu 387
QY 1393 TTGAATTTCTCAGTTTTCAGGCTGGCTTTTATTGGGCTGGCAGTTCCTGGGCTGATT 1452
Db 388 IleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle 407

QY 1453 TATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGGTGCCACTGTTTCATCCCA 1512
Db 408 ValLeuArgTrpLysLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
QY 1513 GCTTTGTTTCTTCACATGCCCTCTTCATGGTGGCTTCCCTCTATTCGGACCCATTT 1572
Db 428 IleIleTyrLeuLeuPheTrpAlaPheLeuLeuValPheSerLeuTrpSerGluProVal 447
QY 1573 AGTACAGGATTTGGTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTTTATT 1632
Db 448 ValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPheLeuGlyVal 467
QY 1633 ATATGGACAAAGAAACCCAGGTGTTAGATAATAATGTACAGAGAAAATAACAGAACATTA 1692
Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuLeuThrLeuValSer 487
QY 1693 CAATAAATACTGGAAGTTGTA---CCAGAAGAAGATAAG 1728
Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500
RESULT 14
AAAY71061
ID AAY71061 standard; Protein; 535 AA.
XX
AC AAY71061;
XX
DT 29-AUG-2000 (first entry)
XX
DE Human membrane transport protein, MTRP-6.
XX
KW Human; membrane transport protein; MTRP-6; antiinflammatory; cytostatic;
KW antithyroid; immunosuppressive; thyromimetic; antidiabetic; nootropic;
KW antidiarrheic; neuroprotective; antidepressant; nephrotropic; virucide;
KW antihelminthic; protozoacide; antibacterial; neuroleptic; antigout;
KW diagnosis; prevention; treatment; membrane transport disorder; epilepsy;
KW Menkes disease; diabetes; depression; Parkinson's disease; neurological disorder;
KW Alzheimer's disease; depression; schizophrenia; immune disorder; allergy;
KW inflammatory disorder; Hashimoto's thyroiditis; microbial infection; cancer;
KW cell proliferative disorder.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 11
FT /note= "Phosphorylation site"
FT Modified-site 61
FT /note= "Phosphorylation site"
FT Modified-site 103
FT /note= "Phosphorylation site"
FT Modified-site 109
FT /note= "Phosphorylation site"
FT Modified-site 179
FT /note= "Phosphorylation site"
FT Modified-site 292
FT /note= "Phosphorylation site"
FT Modified-site 305
FT /note= "Phosphorylation site"
FT Modified-site 337
FT /note= "Phosphorylation site"
FT Modified-site 363
FT /note= "Phosphorylation site"
FT Modified-site 487
FT /note= "Phosphorylation site"
FT Modified-site 502
FT /note= "Phosphorylation site"
FT Region 36..474
FT /note= "Amino acid permeases signature sequence"
FT Domain 76..96
FT /label= Transmembrane_domain
FT Domain 111..131
FT /label= Transmembrane_domain

Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrProIleProAla 367
QY 1333 GTATTGTTTGCACCCCTTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTT 1392
Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu 387
QY 1393 TTGAATTTCCCTCAGTTTTCGCCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATT 1452
Db 388 IleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle 407
QY 1453 TATCTCGATACAAATGCCAGATATGCATCGTCCTTTCAAGGTGCCACTGTTTCATCCCA 1512
Db 408 ValLeuArgTyrLysLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
QY 1513 GCTTTGTTTTCCTTCACATGCCCTCTTTCATGGTTGCCCTTTCCCTCTATTTCGGACCCATT 1572
Db 428 IleIleTyrLeuLeuPheTyrAlaPheLeuLeuValPheSerLeuTyrSerGluProVal 447
QY 1573 AGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATT 1632
Db 448 ValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPheLeuGlyVal 467
QY 1633 ATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTCAGAGAAAATAACCCAGAACATTA 1692
Db 468 TyrTyrGlnHisLysProLysCysPheSerAspPheIleGluLeuLeuThrLeuValSer 487
QY 1693 CAAATAATACTGGAAGTTGTA---CCAGAGAAGATAAG 1728
Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500

Search completed: April 16, 2003, 16:09:32
Job time : 86 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 09:34:04 ; Search time 6504 Seconds
(without alignments)
10018.632 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: { OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2239	100.0	2239	6	AX321909	AX321909 Sequence
2	1867	83.4	5981	6	AX321910	AX321910 Sequence
3	1816	81.1	3144	9	HSA277882	AJ277882 Homo sapi
4	1722	76.9	2482	9	AF252872	AF252872 Homo sapi
5	1704	76.1	2155	9	BC012087	BC012087 Homo sapi
6	1459	65.2	1874	9	AF200708	AF200708 Homo sapi
7	1454	64.9	1861	9	AB026891	AB026891 Homo sapi
8	1339	59.8	2000	9	AB040875	AB040875 Homo sapi
9	479	21.4	64196	9	AC110804	AC110804 Homo sapi
10	281	12.6	337	6	AX321911	AX321911 Sequence
11	269	12.0	596	6	AX351341	AX351341 Sequence
12	182	8.1	161280	9	AC093903	AC093903 Homo sapi
13	173	7.7	1683	9	AB042201	AB042201 Homo sapi
14	128	5.7	3994	9	AC116610	AC116610 Homo sapi
15	107	4.8	575	6	AX341060	AX341060 Sequence
16	84	3.8	632	11	G61322	G61322 SHGC-85475
17	52	2.3	492	10	AB037653	AB037653 Mus muscu
18	52	2.3	2216	10	AB022345	AB022345 Mus muscu
19	52	2.3	183949	2	AC118181	AC118181 Rattus no
20	49	2.2	724	9	HSNC2ALPH	X96506 H.sapiens m
21	49	2.2	788	9	BC010025	BC010025 Homo sapi
22	45	2.0	983	6	AX011754	AX011754 Sequence
23	45	2.0	92203	2	AP001920	AP001920 Homo sapi
24	45	2.0	162511	9	AC009470	AC009470 Homo sapi
25	45	2.0	182547	9	AP001201	AP001201 Homo sapi
26	39	1.7	942	10	AB037652	AB037652 Mus muscu
27	37	1.7	97773	2	AC106406	AC106406 Rattus no
28	37	1.7	183949	2	AC118181	AC118181 Rattus no
29	35	1.6	40690	9	AC112186	AC112186 Homo sapi
30	35	1.6	70638	9	AC079306	AC079306 Homo sapi
31	35	1.6	110000	2	AC023632_0	AC023632 Homo sapi
32	35	1.6	110000	2	AC023632_2	Continuation (3 Of
33	35	1.6	143183	2	AC026927	AC026927 Homo sapi
34	35	1.6	145695	2	AC117533	AC117533 Homo sapi
35	35	1.6	159692	2	AC013693	AC013693 Homo sapi
36	35	1.6	161806	2	AC016495	AC016495 Homo sapi
37	35	1.6	195623	2	AP004218	AP004218 Homo sapi
38	34	1.5	149336	2	AF315312	AF315312 Homo sapi
39	34	1.5	159104	9	AC108925	AC108925 Homo sapi
40	33	1.5	33667	9	HSU227D1	Z68908 Human DNA S
41	33	1.5	38041	9	AC002985	AC002985 Human DNA
42	33	1.5	39617	9	AP000348	AP000348 Homo sapi
43	33	1.5	43472	2	AC066605	AC066605 Homo sapi
44	33	1.5	44430	9	AC092309	AC092309 Homo sapi
45	33	1.5	44869	9	AB022537	AB022537 Homo sapi

ALIGNMENTS

RESULT 1
AX321909
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX321909
Sequence 440 from Patent WO0172295.
AX321909
AX321909.1 GI:17906515
human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
Mannion,J. and Kalos,M.D.

TITLE	Compositions and methods for the therapy and diagnosis of lung cancer									
JOURNAL	Patent: WO 0172295-A 440 04-OCT-2001; CORIXA CORPORATION (US)									
FEATURES	Location/Qualifiers									
source	1. .2239 /organism="Homo sapiens" /db_xref="taxon:9606"									
BASE COUNT	619 a	444 c	493 g	683 t						
ORIGIN										
Query Match	100.0%;		Score 2239;	DB 6;	Length 2239;					
Best Local Similarity	100.0%;		Pred. No. 0;							
Matches 2239;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;		
QY	1	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCTGGGTGACAGTGAGACTCTGTCTCAAAAC	60							
Db	1	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCTGGGTGACAGTGAGACTCTGTCTCAAAAC	60							
QY	61	AGAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGAGAAATTCAGGCCAATTTGTGG	120							
Db	61	AGAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGAGAGAAATTCAGGCCAATTTGTGG	120							
QY	121	CATAGATTTTATCATATCTGGATTTTGTGGATTTCTTTTCTCATCACTGGATTCA	180							
Db	121	CATAGATTTTATCATATCTGGATTTTGTGGATTTCTTTTCTCATCACTGGATTCA	180							
QY	181	GGAAGCCTGTTGTGTCCACCAATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGG	240							
Db	181	GGAAGCCTGTTGTGTCCACCAATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGG	240							
QY	241	AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGSCAGGACGCCCTTTTCAGGAAGAGAC	300							
Db	241	AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGSCAGGACGCCCTTTTCAGGAAGAGAC	300							
QY	301	GCCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGACAGCTGAAGAGGAAAGTC	360							
Db	301	GCCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGACAGCTGAAGAGGAAAGTC	360							
QY	361	ACTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCACATCATTTGAGCAGGAATCTTCATC	420							
Db	361	ACTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCACATCATTTGAGCAGGAATCTTCATC	420							
QY	421	TCTCCTAAGGGCGTGCTCCAGAACACGGGCAGCGTGGGCATGTCTCTGACCATCTGGACG	480							
Db	421	TCTCCTAAGGGCGTGCTCCAGAACACGGGCAGCGTGGGCATGTCTCTGACCATCTGGACG	480							
QY	481	GTGTGTGGGGTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATATA	540							
Db	481	GTGTGTGGGGTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATATA	540							
QY	541	AAGAAATCTGGAGGTCAATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTT	600							
Db	541	AAGAAATCTGGAGGTCAATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTT	600							
QY	601	GTACGAGTCTGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTG	660							
Db	601	GTACGAGTCTGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTG	660							
QY	661	GCATTTGGACGCTACATCTCGAACCAATTTTATTCAATGTGAAATCCCTGAACCTTGGC	720							
Db	661	GCATTTGGACGCTACATCTCGAACCAATTTTATTCAATGTGAAATCCCTGAACCTTGGC	720							
QY	721	ATCAAGCTCATTAACAGTGTGGGCATAAAGTGTAGTGATGGTCTCTAAATAGCATGAGTGTC	780							
Db	721	ATCAAGCTCATTAACAGTGTGGGCATAAAGTGTAGTGATGGTCTCTAAATAGCATGAGTGTC	780							
QY	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTTCGAAGCTCACAGCAATTTCTGATA	840							
Db	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTTCGAAGCTCACAGCAATTTCTGATA	840							
QY	841	ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGACGCC	900							

Db	841	ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGACGCC	900
QY	901	TTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTAT	960
Db	901	TTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTAT	960
QY	961	GCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCCCTGAAAAA	1020
Db	961	GCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCCCTGAAAAA	1020
QY	1021	ACCATTTCCCTTGCATATATGTATATCCATGGCCATTGTACCACTTGGCTATGTGCTGACA	1080
Db	1021	ACCATTTCCCTTGCATATATGTATATCCATGGCCATTGTACCACTTGGCTATGTGCTGACA	1080
QY	1081	AATGTGGCCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCA	1140
Db	1081	AATGTGGCCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCA	1140
QY	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTCCGATCTTTGTGGCC	1200
Db	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTCCGATCTTTGTGGCC	1200
QY	1201	CTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAGGTTATTCTATGTT	1260
Db	1201	CTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAGGTTATTCTATGTT	1260
QY	1261	GGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGGCAAGCACACT	1320
Db	1261	GGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGGCAAGCACACT	1320
QY	1321	CCTCTACCAAGCTGTTATGTTTGCACCTTTGCCAGGTGGCTTTTATTGGGCTGGCAGTT	1380
Db	1321	CCTCTACCAAGCTGTTATGTTTGCACCTTTGCCAGGTGGCTTTTATTGGGCTGGCAGTT	1380
QY	1381	CTCGACAGTCTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTT	1440
Db	1381	CTCGACAGTCTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTT	1440
QY	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGCTTTCAGAGTGCCA	1500
Db	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGCTTTCAGAGTGCCA	1500
QY	1501	CTGTTTCATCCCGCTTGTGTTTCTTCACATGCCCTTTCATGTTGGCTTCCCTCTAT	1560
Db	1501	CTGTTTCATCCCGCTTGTGTTTCTTCACATGCCCTTTCATGTTGGCTTCCCTCTAT	1560
QY	1561	TCGGACCCCATTTAGTACAGGGATTGGCTTCTGTCATCACTCTGACTGGAGTCCCTGCGTAT	1620
Db	1561	TCGGACCCCATTTAGTACAGGGATTGGCTTCTGTCATCACTCTGACTGGAGTCCCTGCGTAT	1620
QY	1621	TATCTCTTTATTTATGGGACAAAGAAACCCAGGTGTTTGAATAATGTCAGAGAAAAATA	1680
Db	1621	TATCTCTTTATTTATGGGACAAAGAAACCCAGGTGTTTGAATAATGTCAGAGAAAAATA	1680
QY	1681	ACCAGAACATTTACAAATAACTGGAAGTTGTACCAGAAAGAGATAAGTTATGAACATAAT	1740
Db	1681	ACCAGAACATTTACAAATAACTGGAAGTTGTACCAGAAAGAGATAAGTTATGAACATAAT	1740
QY	1741	GGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAT	1800
Db	1741	GGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAT	1800
QY	1801	TTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGATAAAACAAAGGAGTCAGTTATTTTA	1860
Db	1801	TTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGATAAAACAAAGGAGTCAGTTATTTTA	1860
QY	1861	TTCATATATTTTAGCATATTCGAACATAATTTCTAAGAAAAATTTAGTTATACTCTATGTAG	1920
Db	1861	TTCATATATTTTAGCATATTCGAACATAATTTCTAAGAAAAATTTAGTTATACTCTATGTAG	1920
QY	1921	TTATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTCTTGAGTCTCTGATACC	1980
Db	1921	TTATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTCTTGAGTCTCTGATACC	1980

QY 1981 TACCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGGTCAATCTCTACAACAT 2040
Db 1981 TACCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGGTCAATCTCTACAACAT 2040
QY 2041 ATGTTAGCAGCGCAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGTT 2100
Db 2041 ATGTTAGCAGCGCAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGTT 2100
QY 2101 TTGTAAGATGGTTTTACACACTACAGATGTCTATAGTGTGAAAAGTGTCTCAATTTCTG 2160
Db 2101 TTGTAAGATGGTTTTACACACTACAGATGTCTATAGTGTGAAAAGTGTCTCAATTTCTG 2160
QY 2161 AAAAAAGCATACATCATGATTATGGCAAGAGGAGAGAAAGGTAGAGCTGTCTTAAATT 2220
Db 2161 AAAAAAGCATACATCATGATTATGGCAAGAGGAGAGAAAGGTAGAGCTGTCTTAAATT 2220
QY 2221 TATTAATAAAAAA 2239
Db 2221 TATTAATAAAAAA 2239
RESULT 2
AX321910
LOCUS AX321910 5981 bp DNA linear PAT 15-DEC-2001
DEFINITION Sequence 441 from Patent WO0172295.
ACCESSION AX321910
VERSION AX321910.1 GI:17906518
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Read,S.G., Lodes,M.J., Mohamath,R., Secret,H., Benson,D.R.,
Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
Mannion,J. and Kalos,M.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0172295-A 441 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source 1. 5981
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1858 a 1029 c 1098 g 1996 t
ORIGIN
Query Match 83.4%; Score 1867; DB 6; Length 5981;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 278 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 337
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 453
Db 338 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 397
QY 454 GTGGGCATGTCTGTGACCATCTGGACGGTGTGTGGGGTCTGTCACTATTTGGAGCTTTG 513
Db 398 GTGGGCATGTCTGTGACCATCTGGACGGTGTGTGGGGTCTGTCACTATTTGGAGCTTTG 457
QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCAATACACATATATTTTG 573
Db 458 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCAATACACATATATTTTG 517
QY 574 GAAGTCTTTGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATAATACGC 633
Db 518 GAAGTCTTTGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATAATACGC 577
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAAACCATTTTTT 693

Db 578 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTTTT 637
QY 694 ATTCAATGTGAATCCCTGAACTTGGATCAAGCTCATTACAGCTGTGGGCATAAAGTGA 753
Db 638 ATTCAATGTGAATCCCTGAACTTGGATCAAGCTCATTACAGCTGTGGGCATAAAGTGA 697
QY 754 GTGATGGTCCCTAAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTCTTAACC 813
Db 698 GTGATGGTCCCTAAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTCTTAACC 757
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGAGCTAATATAA 873
Db 758 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGAGCTAATATAA 817
QY 874 GGTCAACGCAGAACTTTAAAGACGCCCTTTTCAAGGAAGAGATTCAAGTATTACGCGGTG 933
Db 818 GGTCAACGCAGAACTTTAAAGACGCCCTTTTCAAGGAAGAGATTCAAGTATTACGCGGTG 877
QY 934 CCACTGGCTTTTATATATGGAATGTATGATATGCTGGCTGGTGTACCTCAACTTTGTT 993
Db 878 CCACTGGCTTTTATATATGGAATGTATGATATGCTGGCTGGTGTACCTCAACTTTGTT 937
QY 994 ACTGAAGAAGTAGAAAAACCTTGAAAAAACCATTTCCCTTGCATATATATATCCATGGCC 1053
Db 938 ACTGAAGAAGTAGAAAAACCTTGAAAAAACCATTTCCCTTGCATATATATATCCATGGCC 997
QY 1054 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113
Db 998 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1057
QY 1114 GAGCTGCTCTCAAAATGCAGTGGGAGTGACCTTTTCTGAGCGGCTACTTCCAGAAATTC 1173
Db 1058 GAGCTGCTCTCAAAATGCAGTGGGAGTGACCTTTTCTGAGCGGCTACTTCCAGAAATTC 1117
QY 1174 TCATTAGCAGTCCGATCTTTTGTGGCTCTCTGCTTTGGCTCCATGAACGGTGGTGTG 1233
Db 1118 TCATTAGCAGTCCGATCTTTTGTGGCTCTCTGCTTTGGCTCCATGAACGGTGGTGTG 1177
QY 1234 TTTGCTGCTCCAGGTTATTTCTATGTTGCTGCTCGAGAGGGTCACTTCCAGAAATCCTC 1293
Db 1178 TTTGCTGCTCCAGGTTATTTCTATGTTGCTGCTCGAGAGGGTCACTTCCAGAAATCCTC 1237
QY 1294 TCCATGATTATGTCGCAAGCACACTCTCTACAGCTGTTATTTGTTTGCACCTTTG 1353
Db 1238 TCCATGATTATGTCGCAAGCACACTCTCTACAGCTGTTATTTGTTTGCACCTTTG 1297
QY 1354 ACAATGATAAATGCTCTCTCTGGAGACCTCGACAGTCTTTGAATTTCCCTCAGTTTGGC 1413
Db 1298 ACAATGATAAATGCTCTCTCTGGAGACCTCGACAGTCTTTGAATTTCCCTCAGTTTGGC 1357
QY 1414 AGGTGGCTTTTATTGGGCTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 1358 AGGTGGCTTTTATTGGGCTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1417
QY 1474 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTCCTTCACATGC 1533
Db 1418 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTCCTTCACATGC 1477
QY 1534 CTCTTCATGGTTGCCCTTTCCCTCTATTTCGGACCATTTAGTACAGGATTTGGCTTCGTC 1593
Db 1478 CTCTTCATGGTTGCCCTTTCCCTCTATTTCGGACCATTTAGTACAGGATTTGGCTTCGTC 1537
QY 1594 ATCACTCTGACTGGAGTCCCTCGCTATTTATCTTTATATATGGGACAAAGAACCCAGG 1653
Db 1538 ATCACTCTGACTGGAGTCCCTCGCTATTTATCTTTATATATGGGACAAAGAACCCAGG 1597
QY 1654 TGGTTTGAATAATGTACAGAGAAAAATAACAGAACATTTACAAATAATACTGGAAGTTGTA 1713
Db 1598 TGGTTTGAATAATGTACAGAGAAAAATAACAGAACATTTACAAATAATACTGGAAGTTGTA 1657
QY 1714 CCAGAAAGAGATAAGTTATGAACCTAATGGACTTGAGATCTTTGGCAATCTGCCCAAGGGGA 1773

Db 1658 CCAGAGAAGATAAGTTATGAACATAAGTATGAACCTAGGACTTGAGATCTTGGCAATCTGCCCAAGGGGA 1717

QY 1774 GACACAAAATAGGGATTTTACTTCATTTTCTGAAAGCTTAGAGAAATTACAACCTTTGGTG 1833

Db 1718 GACACAAAATAGGGATTTTACTTCATTTTCTGAAAGCTTAGAGAAATTACAACCTTTGGTG 1777

QY 1834 ATAAACAAAGGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCT 1893

Db 1778 ATAAACAAAGGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCT 1837

QY 1894 AAGAAATTTAGTTATAACTCTATGTAGTTATAGAAAGTGAATATGCAGTTATTCATGAG 1953

Db 1838 AAGAAATTTAGTTATAACTCTATGTAGTTATAGAAAGTGAATATGCAGTTATTCATGAG 1897

QY 1954 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAAGACTAGACAA 2013

Db 1898 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAAGACTAGACAA 1957

QY 2014 TTACTATGTGTCATCTCTACACATATGTTAGCACGGGCAAAAGACCTTCAAAATTGAAG 2073

Db 1958 TTACTATGTGTCATCTCTACACATATGTTAGCACGGGCAAAAGACCTTCAAAATTGAAG 2017

QY 2074 ACTGAGATTTTCTGTATATATGGGTTTGTAAAGATGGTTTACACACTACAGATGCT 2133

Db 2018 ACTGAGATTTTCTGTATATATGGGTTTGTAAAGATGGTTTACACACTACAGATGCT 2077

QY 2134 ATACTGTGAAAAGTGTTTTCAATCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 2193

Db 2078 ATACTGTGAAAAGTGTTTTCAATCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 2137

QY 2194 GAGAGAA 2200

Db 2138 GAGAGAA 2144

RESULT 3

HSA277882

LOCUS HSA277882 3144 bp mRNA linear PRI 04-JAN-2002

DEFINITION Homo sapiens mRNA for cystine/glutamate transporter (XCT gene).

ACCESSION AJ277882

VERSION AJ277882.1 GI:18073361

KEYWORDS cystine/glutamate transporter; xct gene.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Borsani,G., Manzoni,M., Palacin,M., Pineda,M. and Gasol,E.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3144)

AUTHORS Bassi,M.T.

TITLE Direct Submission

JOURNAL Submitted (11-MAY-2000) Bassi M.T., Telethon Institute of Genetics and Medicine, Via Olgettina 58, 20132 Milan, ITALY

FEATURES

source

1. .3144

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="4"

/map="4q28-q32"

/cell_line="NT2, undifferentiated teratocarcinoma cell line"

1. .3128

/gene="XCT"

1. .1506

/gene="XCT"

/codon_start=1

/product="cystine/glutamate transporter"

/protein_id="CAC81905.1"

/db_xref="GI:18073362"

/translation="MVRKPVVSTISKGYLQGNVNGRLPSLKNKEPPGQEKVQLRKV TLLRGVSIIGTIIGAGIFISPKGVLQNTGSGMSLTITWVGVLSLFGALSVAELGT TIKKSGGHYTIIEVFGPLPAFVRVWVELLIIRPAATAVISLAFGRYILEPFFIQCEI

PELAIKLITAVGITVVMVLSMSVSWSARIOIQLFTFCKLTAIIIIIVPGVMQLIKGQT QNEKDAFSGRDSITRLPLAFYYGMYAYAGWFYLNFTVEEVENKFTIPLAICISMAI VTIGYVLTNVAYFTTINAELELLSNVAVTFSERLVGNFSLAVPIFVALSCFGSMNGG VFAVSRLFYVASREGLHPEILSMIHVRKHPLPAVIVLHPLTMIMLPSGDLDSLNLFL SFARWLFGLAVAGLIYLRKCPDPMHRPFKVLPIPALFSFTCLFVALSLYSDPFT GIGFVITLTGVPAYVLFIIWDKKPRWRFRIMSEKITRTLIILEVVPEDKL"

repeat_region 2857. .3127 /rpt_family="Sine/Alu"

polyA_site 3128 /gene="XCT"

BASE COUNT 912 a 602 c 649 g 981 t

ORIGIN

Query Match 81.1%; Score 1816; DB 9; Length 3144;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGC 393

Db 106 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGC 165

QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 453

Db 166 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 225

QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTGAGCTTTG 513

Db 226 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTGAGCTTTG 285

QY 514 TCTTATGCTGAATTTGGAAACAACATAAAGAAATCTGGAGGTCATTACACATATATTTG 573

Db 286 TCTTATGCTGAATTTGGAAACAACATAAAGAAATCTGGAGGTCATTACACATATATTTG 345

QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATAATACGC 633

Db 346 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATAATACGC 405

QY 634 CCTGCAGCTACTGTCTGATATATCCCTGGCATTTTGACGCTACATTTCTGGAACCATTTT 693

Db 406 CCTGCAGCTACTGTCTGATATATCCCTGGCATTTTGACGCTACATTTCTGGAACCATTTT 465

QY 694 ATTCAATGTGAAATCCCTGAACTTCCGATCAAGCTCATTTACAGCTGTGGGCATAACTGTA 753

Db 466 ATTCAATGTGAAATCCCTGAACTTCCGATCAAGCTCATTTACAGCTGTGGGCATAACTGTA 525

QY 754 GTGATGGTCTTAATAGCATGAGTGTGAGTGGAGCGCCGGATCCAGATTTCTTAACC 813

Db 526 GTGATGGTCTTAATAGCATGAGTGTGAGTGGAGCGCCGGATCCAGATTTCTTAACC 585

QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATAATAA 873

Db 586 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATAATAA 645

QY 874 GGTCAAAACGCAGAACTTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGATTTACGGGTTG 933

Db 646 GGTCAAAACGCAGAACTTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGATTTACGGGTTG 705

QY 934 CCAGTGGCTTTTATATGGAATGTATGCATATGCTGGTGGTGTACCTCAACTTTGTT 993

Db 706 CCAGTGGCTTTTATATGGAATGTATGCATATGCTGGTGGTGTACCTCAACTTTGTT 765

QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTCCTTGCATATGTATATATCCATGGCC 1053

Db 766 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTCCTTGCATATGTATATATCCATGGCC 825

QY 1054 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113

Db 826 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 885

QY 1114 GAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGACGGGCTACTGGGAAATTTTC 1173

Db 886 GAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGACGGGCTACTGGGAAATTTTC 945

QY 1174 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTG 1233
Db 946 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTG 1005
QY 1234 TTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATCCTC 1293
Db 1006 TTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATCCTC 1065
QY 1294 TCCATGATTCATGTCGCAAGCACACTCCTCTACCGAGCTGTTATGTTTGGACCCCTTG 1353
Db 1066 TCCATGATTCATGTCGCAAGCACACTCCTCTACCGAGCTGTTATGTTTGGACCCCTTG 1125
QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGCC 1413
Db 1126 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGCC 1185
QY 1414 AGGTGGCTTTTATTTGGGCTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCA 1473
Db 1186 AGGTGGCTTTTATTTGGGCTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCA 1245
QY 1474 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTCCTTCACATGC 1533
Db 1246 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTCCTTCACATGC 1305
QY 1534 CTCTCATGTTGCCCTTTCCCTCTATTTCGGACCCATTAGTACAGGATGGCTTCGTC 1593
Db 1306 CTCTCATGTTGCCCTTTCCCTCTATTTCGGACCCATTAGTACAGGATGGCTTCGTC 1365
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTATTATATGGGACAGAAACCCAGG 1653
Db 1366 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTATTATATGGGACAGAAACCCAGG 1425
QY 1654 TGGTTAGAAATAGTTCAGAGAAAATAACCCAGAACATTACAAATAATCTGGAAGTTGA 1713
Db 1426 TGGTTAGAAATAGTTCAGAGAAAATAACCCAGAACATTACAAATAATCTGGAAGTTGA 1485
QY 1714 CCAGAGAAGATAAGTTATGAACATAATGGACTTCAGATCTTGGCAATCTGCCCAAGGGGA 1773
Db 1486 CCAGAGAAGATAAGTTATGAACATAATGGACTTCAGATCTTGGCAATCTGCCCAAGGGGA 1545
QY 1774 GACACAAAATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTG 1833
Db 1546 GACACAAAATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTG 1605
QY 1834 ATAAACAAAAGGAGTCAGTTATTTTATTTCATATATTTTAGCATATATTCGAACATAATTTCT 1893
Db 1606 ATAAACAAAAGGAGTCAGTTATTTTATTTCATATATTTTAGCATATATTCGAACATAATTTCT 1665
QY 1894 AAGAAATTTAGTTATAACTCTATGTAGTTATAGAAAGTGAATATGCACTTATCTATGAG 1953
Db 1666 AAGAAATTTAGTTATAACTCTATGTAGTTATAGAAAGTGAATATGCACTTATCTATGAG 1725
QY 1954 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAA 2013
Db 1726 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAA 1785
QY 2014 TTACTATGTTGTCATCTCTACAACATATGTTAGCACGGCAAGAACCTTCAAATTTGAAG 2073
Db 1786 TTACTATGTTGTCATCTCTACAACATATGTTAGCACGGCAAGAACCTTCAAATTTGAAG 1845
QY 2074 ACTGAGATTTTCTGTATATATGGGTTTGTAAAGATGGTTTTTACACACTACAGATGTCT 2133
Db 1846 ACTGAGATTTTCTGTATATATGGGTTTGTAAAGATGGTTTTTACACACTACAGATGTCT 1905
QY 2134 ATACTGTGAAAAGTGTTTTCAATTTCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 2193
Db 1906 ATACTGTGAAAAGTGTTTTCAATTTCTGAAAAAAGCATACATCATGATTATGGCAAAGAG 1965
QY 2194 GAGAGAA 2200
Db 1966 GAGAGAA 1972

RESULT 4
AF252872
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AF252872 2482 bp mRNA linear PRI 02-MAY-2001
Homo sapiens cystine/glutamate transporter xCT mRNA, complete cds.
AF252872.1 GI:13924719
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2482)
Chancy,C.D., Kekuda,R., Wang,H., Huang,W., Prasad,P.D., Smith,S.B.
and Ganapathy,V.
Structure, Function and Regulation of Human Cystine/Glutamate
Transporter in Retinal Pigment Epithelial Cells
Unpublished
2 (bases 1 to 2482)
Wang,H., Prasad,P.D. and Ganapathy,V.
Direct Submission
Submitted (05-APR-2000) Biochemistry & Molecular Biology, Medical
College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
Location/Qualifiers
1. .2482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
232..1737
/codon_start=1
/product="cystine/glutamate transporter xCT"
/protein_id="AAK49111.1"
/translation="MVRKPVVSTISKGYLOGNVNRLPSLGNKEPPGQEKVQLKRKV
TLRGVSIITIGTIGAGIFISPKGVQLQNTSGVMSLTITWVGVLSLFGALSVAELGT
TIKSGGHYTYILEVFGPLPAFVRVWVELLIIRPAATAVISLAFGRVILEPFFIQCEI
PELAIKLITAVGITVVMVNSMSVSARSQIFLTCKLTAILIIVPGVMQLIKGQT
QNFKDAFSGRDSITRLPLAFYGYMYAYAGWFYLNFTVEEVENPEKTIPLAICISMAI
VTIGYVLTNVAFTTINAEELLSNAVAVTFSERLLGNFSLAVPIFVALSCFGSMNGG
VFVSRIFYVASREGHLPEILSMIHVRKHTPLPAVIVLHPLTMIMLPSGDLDSLLNFL
SFARWLFGLAVAGLIYLRKCPDPMHRPKVPLFIPALFSTCLEMVALSLSDPFSF
GIGFVITLTGVPAYVLFIIWDKKPRWFRIMSEKITRTLQIILEVVPEDKL"

BASE COUNT 735 a 484 c 521 g 742 t
ORIGIN

Query Match 76.9%; Score 1722; DB 9; Length 2482;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1772; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGAGTCTCCATTATATTGGC 393
Db 337 GAGAAAGTGCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGAGTCTCCATTATCATTTGGC 396
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTGGGTCCTGTCACTATTTGGAGCTTG 453
Db 397 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTGGGTCCTGTCACTATTTGGAGCTTG 456
QY 454 GTGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTTGGAGCTTG 513
Db 457 GTGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTTGGAGCTTG 516
QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCATTACACATATATTTTG 573
Db 517 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCATTACACATATATTTTG 576
QY 574 GAAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACTCCCTCATAATACGC 633
Db 577 GAAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACTCCCTCATAATACGC 636
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGACGCTACATTTCTGGACCATTTTTT 693
Db 637 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGACGCTACATTTCTGGACCATTTTTT 696
QY 694 ATTCAATGTGAAATCCCTGAACTTGGCATCAAGCTCAAGCTGTGGGCATAACTGTA 753

CDS

/lab_host="DH10B-R"
/note="Vector: pOTB7"
161. .1666
/codon_start=1
/product="Similar to solute carrier family 7, (cationic
amino acid transporter, y+ system) member 11"
/protein_id="AAH12087.1"
/db_xref="GI:15082352"
/translation="MVRKPVSTISKGGYLOGNVNRLPSLGNKEPPGQEKVQLRKRV
TLRGVSIIGTIIGAGIFISPKGLQNTSGVMSLTIWTVGVLSLFGLSVAELGT
TIKKSQGHYTYILEVFGPLPAFVRVWVLELLIIRPAATAVISLAFGRYILEPFIQCEI
PELAIKLITAVGITVVMVLSMSVSWSARIOIFLTFCKLITAILIIVPGVMQLIKGQT
QNFKDAFSGRDSSITRLPLAFYGYMYAGWFYLFNFVEEVENPEKTIPLAICISMAI
VTIGVLTNVAYFTTINAEELLSNAVAVTFSERLGNFSLAVPIFVALSCFSGMNGG
VFAVSRLEYVASREGHLPILSMIHVRKHTPLPAVIVLHPLTMIMLFSGLDLSLINF
SFARWLFGLAVAGLIYLRVKCPDMHRPFKVPPLFIPALFSTCLFMVALSLYSDPFST
GIGFVITLTGVPAYYLFIIWKKPRWFRIMSEKITRLQLIILEVVPEEDKL"

BASE COUNT 599 a 443 c 452 g 661 t
ORIGIN

Query Match		76.1%;	Score 1704;	DB 9;	Length 2155;
Best Local Similarity		99.8%;	Pred. No. 0;		
Matches 1854;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	334	GAGAAAGTGCAGCTGAAGAGGAAAAGTCATTACTGAGGGGAGTCTCCATTATCATTTGGC	393		
Db	266	GAGAAAGTGCAGCTGAAGAGGAAAAGTCATTACTGAGGGGAGTCTCCATTATCATTTGGC	325		
QY	394	ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTTAAGGGCGTGCTCCAGAACACGGGCAGC	453		
Db	326	ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTTAAGGGCGTGCTCCAGAACACGGGCAGC	385		
QY	454	GTGGGCATGTCTCTGACCATCTGGACCGGTGTGGGGTCTGTCTGCTACTATTGGAGCTTTG	513		
Db	386	GTGGGCATGTCTCTGACCATCTGGACCGGTGTGGGGTCTGTCTGCTACTATTGGAGCTTTG	445		
QY	514	TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTAACACATATATTTTG	573		
Db	446	TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTAACACATATATTTTG	505		
QY	574	GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGGAACCTCTCATATAACGC	633		
Db	506	GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGGAACCTCTCATATAACGC	565		
QY	634	CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATCTGGAACCATTTTGT	693		
Db	566	CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATCTGGAACCATTTTGT	625		
QY	694	ATTCAATGTGAATCCCTGAACCTTGGATCAAGCTCATTACAGCTGTGGCATAAAGTGA	753		
Db	626	ATTCAATGTGAATCCCTGAACCTTGGATCAAGCTCATTACAGCTGTGGCATAAAGTGA	685		
QY	754	GTGATGGTCTTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACC	813		
Db	686	GTGATGGTCTTAAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACC	745		
QY	814	TTTGTCAAGCTCACAGCAATCTGATAAATTATAGTCCCTGGAGTTATGCAGCTAATTAA	873		
Db	746	TTTGTCAAGCTCACAGCAATCTGATAAATTATAGTCCCTGGAGTTATGCAGCTAATTAA	805		
QY	874	GGTCAACCGCAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGGCGTTG	933		
Db	806	GGTCAACCGCAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGGCGTTG	865		
QY	934	CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT	993		
Db	866	CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT	925		
QY	994	ACTGAAGAGTAGAAAACCCCTGAAAAACCATTCCTTTCGCAATATGTATATCCATGGCC	1053		
Db	926	ACTGAAGAGTAGAAAACCCCTGAAAAACCATTCCTTTCGCAATATGTATATCCATGGCC	985		
QY	1054	ATTGTACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG	1113		

Db	986	ATTTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG	1045		
QY	1114	GAGCTGCTGCTTTCAAAATGCAGTGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC	1173		
Db	1046	GAGCTGCTGCTTTCAAAATGCAGTGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC	1105		
QY	1174	TCATTAGCAGTTCAGATCTTTTGTGCCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTG	1233		
Db	1106	TCATTAGCAGTTCAGATCTTTTGTGCCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTG	1165		
QY	1234	TTTGTGCTCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATCCTC	1293		
Db	1166	TTTGTGCTCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATCCTC	1225		
QY	1294	TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTTGTTGCACCTTTG	1353		
Db	1226	TCCATGATTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTTGTTGCACCTTTG	1285		
QY	1354	ACAATGATAATGCTCTTCTGAGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGC	1413		
Db	1286	ACAATGATAATGCTCTTCTGAGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGC	1345		
QY	1414	AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA	1473		
Db	1346	AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA	1405		
QY	1474	GATATGCATCGTCTTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTCCTTCACATGC	1533		
Db	1406	GATATGCATCGTCTTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTCCTTCACATGC	1465		
QY	1534	CTCTTCATGCTTGGCCCTTTCCCTCTATTTCGGACCCCATTTAGTACAGGATTGGCTCGTC	1593		
Db	1466	CTCTTCATGCTTGGCCCTTTCCCTCTATTTCGGACCCCATTTAGTACAGGATTGGCTCGTC	1525		
QY	1594	ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGGACAAAGAACCCAGG	1653		
Db	1526	ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGGACAAAGAACCCAGG	1585		
QY	1654	TGGTTTGAATAATGTGACAGAAAAATAACCAGAACATTAACAAATAATCTGGAAGTTGTA	1713		
Db	1586	TGGTTTGAATAATGTGACAGAAAAATAACCAGAACATTAACAAATAATCTGGAAGTTGTA	1645		
QY	1714	CCAGAAAGAGATAAGTTATGAACATAATGGACTTGAGATCTTGCAATCTGCCAAGGGGA	1773		
Db	1646	CCAGAAAGAGATAAGTTATGAACATAATGGACTTGAGATCTTGCAATCTGCCAAGGGGA	1705		
QY	1774	GACACAAAATAGGGATTTTACTTTCATTTTCTGAAAGTCTAGAGAAATACAACTTTGGTG	1833		
Db	1706	GACACAAAATAGGGATTTTACTTTCATTTTCTGAAAGTCTAGAGAAATACAACTTTGGTG	1765		
QY	1834	ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTAGCATATTTTCGAACTAATTTCT	1893		
Db	1766	ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTATCATATTTTTCGAACTAATTTCT	1825		
QY	1894	AAGAAATTTAGTTATAAATCTATGTAGTTATAGAAAGTGAATATGCAGTTATTTCTATGAG	1953		
Db	1826	AAGAAATTTAGTTATAAATCTATGTAGTTATAGAAAGTGAATATGCAGTTATTTCTATGAT	1885		
QY	1954	TCGCACAATTTCTTGAGTCTCTGATACCTACCTATTGGGGTTAGGAGAAAAGACTAGACAA	2013		
Db	1886	TCGCACAATTTCTTGAGTCTCTGATACCTACCTATTGGGGTTAGGAGAAAAGACTAGACAA	1945		
QY	2014	TTACTATGTGGTCAATCTCTACAACATATGTTAGCACGGCAAGAACTTCAAAATTGAAG	2073		
Db	1946	TTACTATGTGGTCAATCTCTACAACATATGTTAGCACGGCAAGAACTTCAAAATTGAAG	2005		
QY	2074	ACTGAGATTTTCTGTATATATGGGTTTTGTAAAGATGGTTTTACACACTACAGATGTCT	2133		
Db	2006	ACTGAGATTTTCTGTATATATGGGTTTTGTAAAGATGGTTTTACACACTATAGATGTCT	2065		
QY	2134	ATACTGTGAAAAGTGTCTTCAATTTCTGAAAAAAGCATACATCATGATTTATGGCAA	2190		

Db 2066 ATACTGTGAAAAGTGTTCATTTCTGAAAAAAGCATACATCATGATTATGGCAAA 2122

RESULT 6
AF200708
LOCUS
DEFINITION Homo sapiens calcium channel blocker resistance protein CCBRL mRNA, 1874 bp mRNA linear PRI 01-DEC-2000 complete cds.

ACCESSION AF200708
VERSION AF200708.1 GI:11493651

KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1874)
AUTHORS Conklin,D.S. and Beach,D.H.
TITLE CCBRL, novel CD98 light chain implicated in redox control and calcium signaling

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1874)
AUTHORS Conklin,D.S. and Beach,D.H.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Cold Spring Harbor Labs, 1 Bungtown Rd, Cold Spring Harbor, NY 11724, USA

FEATURES
source
1..1874
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="fibrosarcoma HT1080"
 246..1751
 /note="X-c transporter; CD98 light chain"
 /codon_start=1
 /product="calcium channel blocker resistance protein CCBRL"

CDS
 /protein_id="AAG35592.1"
 /db_xref="GI:11493652"
 /translation="MVRKPVVSTISKGGYLOGNVNRLPSLGNKEPPGQEKVQLKRV
 TLRLGVSIIIGTIIGAGIFISPKGVLQNTSGVMSLTIWTCGVLSLFGALSAYAE
 TIKSGGHYTYILEVFGPLPAFVRVWVVELLIIRPAATAVISLAFGRYILEPFFIQCEI
 PELAIKLITAVGITVMVLSMSVSWRSARIQIFLTFCKLTAIIIIIVPGVMQLIKQT
 QNFKDAFSGRDSITRLPLAFYGYMYAGWYLFNFVEEVENPEKTIPLAICISMAI
 VTIGVLTNVAFTTINAEELLSNAVAVTFSERLLGNFSLAVPIFVALSCFGSMNGG
 VFARSLFYVASREGLPEILSMIHRKHPLPAVIVLHPLTMIMLFSGDLSLLNFL
 SFARWLFGLAVAGLIYLYKCPDMHRPFKVPFLFIPALFSFTCLFMVALSLYSDPFS
 GIGFVITLTGVPAYLYFLIWDKKPRWFRIMSEKITRTLQIIILEVVPBEDKL"

BASE COUNT 490 a 407 c 422 g 555 t

ORIGIN
Query Match 65.2%; Score 1459; DB 9; Length 1874;
Best Local-Similarity 99.9%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTGGC 393
 |||||
Db 351 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTGGC 410
 |||||

QY 394 ACCATCATTTGGACAGGAATCTTCATCTCTCCTAAGGGCGTGTCTCCAGAACACGGGCAGC 453
 |||||

Db 411 ACCATCATTTGGACAGGAATCTTCATCTCTCCTAAGGGCGTGTCTCCAGAACACGGGCAGC 470
 |||||

QY 454 GTGGGCATGTCTGTACCATCTGGACGGTGTGTGGGGTCTCTGTCACTATTTGGAGCTTTG 513
 |||||

Db 471 GTGGGCATGTCTGTACCATCTGGACGGTGTGTGGGGTCTCTGTCACTATTTGGAGCTTTG 530
 |||||

QY 514 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCACTACACATATATTTTG 573
 |||||

Db 531 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCACTACACATATATTTTG 590
 |||||

QY 574 GAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAACGC 633
 |||||

Db 591 GAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAACGC 650
 |||||

QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATTTCTGGAACCAATTTT 693
 |||||

Db 651 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATTTCTGGAACCAATTTT 710
 |||||

QY 694 ATTCAATGTGAAATCCCTGAACCTGGCATCAAGCTCATACAGCTGTGGGCATAACTGTA 753
 |||||

Db 711 ATTCAATGTGAAATCCCTGAACCTGGCATCAAGCTCATACAGCTGTGGGCATAACTGTA 770
 |||||

QY 754 GTGATGGTCTTAAATAGCATGAGTGTGAGCTGGAGCGCCGATCCAGATTTTCTTAACC 813
 |||||

Db 771 GTGATGGTCTTAAATAGCATGAGTGTGAGCTGGAGCGCCGATCCAGATTTTCTTAACC 830
 |||||

QY 814 TTTTGCAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATATAA 873
 |||||

Db 831 TTTTGCAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATATAA 890
 |||||

QY 874 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGATTTACCGGGTTG 933
 |||||

Db 891 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGATTTACCGGGTTG 950
 |||||

QY 934 CCACTGGCTTTTATTATGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 993
 |||||

Db 951 CCACTGGCTTTTATTATGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 1010
 |||||

QY 994 ACTGAAGAAGTAGAAAAACCCCTGAAAAAACCATTCCTCTGCAATATGTATATCCATGGCC 1053
 |||||

Db 1011 ACTGAAGAAGTAGAAAAACCCCTGAAAAAACCATTCCTCTGCAATATGTATATCCATGGCC 1070
 |||||

QY 1054 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTAGACCATTAATGCTGAG 1113
 |||||

Db 1071 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAA 1130
 |||||

QY 1114 GAGCTGCTGCTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
 |||||

Db 1131 GAGCTGCTGCTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1190
 |||||

QY 1174 TCATTAGCAGTTCGGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTG 1233
 |||||

Db 1191 TCATTAGCAGTTCGGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTG 1250
 |||||

QY 1234 TTTGCTGTCTCCAGTTTATCTATGTTGGCTCTCGAGAGGGTCACTTCCAGAAATCCCTC 1293
 |||||

Db 1251 TTTGCTGTCTCCAGTTTATCTATGTTGGCTCTCGAGAGGGTCACTTCCAGAAATCCCTC 1310
 |||||

QY 1294 TCCATGATTCATGTCGCAAGCACACTCTCTACAGCTGTTATGTTTGCACCCCTTTG 1353
 |||||

Db 1311 TCCATGATTCATGTCGCAAGCACACTCTCTACAGCTGTTATGTTTGCACCCCTTTG 1370
 |||||

QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1413
 |||||

Db 1371 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1430
 |||||

QY 1414 AGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473
 |||||

Db 1431 AGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1490
 |||||

QY 1474 GATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATCCCAGCTTTGTTTCTTCCATGC 1533
 |||||

Db 1491 GATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATCCCAGCTTTGTTTCTTCCATGC 1550
 |||||

QY 1534 CTCTTCATGTTGGCTTTCCCTCTATTTCGAGCCCATTTAGTACAGGATTTGGCTTCGTC 1593
 |||||

Db 1551 CTCTTCATGTTGGCTTTCCCTCTATTTCGAGCCCATTTAGTACAGGATTTGGCTTCGTC 1610
 |||||

QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTTATATATGGGACAAAGAACCCAGG 1653
 |||||

Db 1611 ATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTTATATATGGGACAAAGAACCCAGG 1670
 |||||

QY 1654 TGGTTTAGAATAATGTCAGAGAAAAATAACAGAACATTAACAAATAACTGGAAGTTGTA 1713
 |||||

Db 1671 TGGTTTAGAATAATGTCAGAGAAAAATAACAGAACATTAACAAATAACTGGAAGTTGTA 1730
 |||||

QY 1714 CCAGAAGAAGATAAGTTATGAACATAATGGACTGAGATCTTTGGCAATCTGCCCAAGGGGA 1773
 |||||

Db 1731 CCAGAGAGAGATAAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCAAGGGGA 1790
QY 1774 GACACAAAATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTG 1833
Db 1791 GACACAAAATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTG 1850
QY 1834 ATAAACAAA 1843
Db 1851 ATAAACAAA 1860

RESULT 7
AB026891
LOCUS AB026891 1861 bp mRNA linear PRI 10-FEB-2001
DEFINITION Homo sapiens mRNA for cystine/glutamate transporter, complete cds.
ACCESSION AB026891
VERSION AB026891.1 GI:5668544
KEYWORDS cystine/glutamate transporter.
SOURCE Homo sapiens fibroblast cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
REFERENCE Sato,H., Tamba,M., Kuriyama-Matsumura,K., Okuno,S. and Bannai,S.
AUTHORS Molecular cloning and expression of human xCT, the light chain of
TITLE amino acid transport system xc-
JOURNAL Antioxidants and Redox Signaling 2, 665-671 (2000)
REFERENCE 2 (bases 1 to 1861)
AUTHORS Sato,H. and Bannai,S.
TITLE Direct Submission
JOURNAL Submitted (30-APR-1999) Hideyo Sato, University of Tsukuba,
Institute of Basic Medical Sciences; Tennodai 1-1-1, Tsukuba,
Ibaraki 305-8575, Japan (E-mail:hideyo-semd.tsukuba.ac.jp,
Tel:81-298-53-3282, Fax:81-298-53-3039)
FEATURES
source Location/Qualifiers
1..1861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="fibroblast"
1..1861
/gene="xCT"
236..1741
/gene="xCT"
/codon_start=1
/product="cystine/glutamate transporter"
/protein_id="BAA82628.1"
/db_xref="GI:5668545"
/translation="MVRKPVVSTISKGYLQNVNRLPSLGNKEPPGQEKVQLRKV
TLRGVSIIGTIIGAGIFISPKVLQNTGVSMSLTIWTCGVLSLFGALSAYELGT
TIKSGGHYTIILEVFGPLPAFVRVWVELLIIRPAATAVISLAFGRYILEPFFIQCEI
PELAIKLITAVGITVVMVNSMSVSWARIQIFLTFCKLTAIIIIYVGMQLIKQT
QNFDAFSGRDSITRLPLAFYGYMYAGWFLNFVTEVENPEKTIPLAICISMAI
VTIGYVLTNAVYFTTINAEELLSNAVAVTFSERLGNFSLAVPIFVALSCFGSMNGG
VFVSRFLFYVASREGHLEILSMIHVRKHTPLPAVIVLHPLTMIMLFSGDLSLLNFL
SFARWLFGLAVAGLIYLRKCPDMHRPFKVPFLFIPALFSFTCLFMVALSLYSDPST
GIGFVITLTGVPAYLFIWDKKPRWFRIMSEKITRTLQIILEVVPEDKL"
BASE COUNT 485 a 402 c 421 g 553 t
ORIGIN

Query Match 64.9%; Score 1454; DB 9; Length 1861;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTCAGCTGAAGAGGAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 341 GAGAAAGTCAGCTGAAGAGGAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 400

QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCTGCTCCAGAACACGGGAGC 453
Db 401 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCTGCTCCAGAACACGGGAGC 460

QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACCTATTTGGAGCTTTG 513

Db 461 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCTCTGACATATTTGGAGCTTTG 520
QY 514 TCTTATGCTGAATTTGGGAACAACACTATAAAGAAATCTGGAGGTCATTACACATATATTTG 573
Db 521 TCTTATGCTGAATTTGGGAACAACACTATAAAGAAATCTGGAGGTCATTACACATATATTTG 580
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAATACGC 633
Db 581 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAATACGC 640
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGGTACATTTGGAACCATTTT 693
Db 641 CCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGGTACATTTGGAACCATTTT 700
QY 694 ATTCAATGTGAAATCCCTGAACCTGCGATCAAGCTCATTACAGCTGTGGGCATAACGTGA 753
Db 701 ATTCAATGTGAAATCCCTGAACCTGCGATCAAGCTCATTACAGCTGTGGGCATAACGTGA 760
QY 754 GTGATGGTCTCTAAATAGCATGATGTGAGTGGAGCGCGGATCCAGATTTTCTTAACC 813
Db 761 GTGATGGTCTCTAAATAGCATGATGTGAGTGGAGCGCGGATCCAGATTTTCTTAACC 820
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGCAGCTAAATAA 873
Db 821 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGCAGCTAAATAA 880
QY 874 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAAGTATTACGGCGTTG 933
Db 881 GGTCAAACGCAGAACTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAAGTATTACGGCGTTG 940
QY 934 CCACTGGCTTTTATATGGAATGTATGCATATGCTGCTGGTGGTTTACCTCAACTTTGTT 993
Db 941 CCACTGGCTTTTATATGGAATGTATGCATATGCTGCTGGTGGTTTACCTCAACTTTGTT 1000
QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCAATTTCCCTTGCATATGTATATCCATGGCC 1053
Db 1001 ACTGAAGAAGTAGAAAACCCCTGAAAAAACCAATTTCCCTTGCATATGTATATCCATGGCC 1060
QY 1054 ATTGTCACCATTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113
Db 1061 ATTGTCACCATTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1120
QY 1114 GAGCTGCTGCTTTCAAATGTCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 1121 GAGCTGCTGCTTTCAAATGTCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTTC 1180
QY 1174 TCATTAGCAGTTCCGATCTTGTGTGCCCTCTCTGCTTGGCTCCATGAACGGTGGTG 1233
Db 1181 TCATTAGCAGTTCCGATCTTGTGTGCCCTCTCTGCTTGGCTCCATGAACGGTGGTG 1240
QY 1234 TTTGCTGTCTCCAGTTATCTATGTTGCGTCTCGAGAGGGTCCACCTTCCAGAAATCCTC 1293
Db 1241 TTTGCTGTCTCCAGTTATCTATGTTGCGTCTCGAGAGGGTCCACCTTCCAGAAATCCTC 1300
QY 1294 TCCATGATTTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTGCACCCCTTTG 1353
Db 1301 TCCATGATTTCATGTCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTGCACCCCTTTG 1360
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC 1413
Db 1361 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC 1420
QY 1414 AGGTGGCTTTTATTTGGCTGGCAGTTGCTGGGTGATTTATCTTCGATACAAATGCCCA 1473
Db 1421 AGGTGGCTTTTATTTGGCTGGCAGTTGCTGGGTGATTTATCTTCGATACAAATGCCCA 1480
QY 1474 GATATGCATCGTCTCTTTCAAGGTGCCACTGTTATCCAGCTTTGTTTCTTCACATGC 1533
Db 1481 GATATGCATCGTCTCTTTCAAGGTGCCACTGTTATCCAGCTTTGTTTCTTCACATGC 1540
QY 1534 CTCTTTCATGGTTGGCTTTCCCTCTATTTCGACCCCATTTTAGTACAGGGATTTGGCTTCGTC 1593

QY 1294 TCCATGATTCATGTCCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTGCACCCCTTG 1353
Db 1201 TCCATGATTCATGTCCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTGCACCCCTTG 1260
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTGAATTCCTCAGTTTGCC 1413
Db 1261 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTGAATTCCTCAGTTTGCC 1320
QY 1414 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATACAAATGCCCA 1473
Db 1321 AGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATACAAATGCCCA 1380
QY 1474 GATATGATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCCCTTCACATGC 1533
Db 1381 GATATGATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCCCTTCACATGC 1440
QY 1534 CTCTTCATGGTGGCTCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATTGGCTTCGTC 1593
Db 1441 CTCTTCATGGTGGCTCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATTGGCTTCGTC 1500
QY 1594 ATCACTCTGACTGGAGTCCCTGGCTATTATCTCTTTATATATGGGACAAAGAACCCAGG 1653
Db 1501 ATCACTCTGACTGGAGTCCCTGGCTATTATCTCTTTATATATGGGACAAAGAACCCAGG 1560
QY 1654 TGGTTTGAATAATGTGAC 1672
Db 1561 TGGTTTGAATAATGTGAC 1579

RESULT 9
AC110804/c
LOCUS AC110804 64196 bp DNA linear PRI 30-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-725C19 from 4, complete sequence.
ACCESSION AC110804
VERSION AC110804.3 GI:19570222
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 64196)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
REFERENCE 2 (bases 1 to 64196)
AUTHORS Pearman, C., Kozlowicz, A. and Dignan, G.
TITLE The sequence of Homo sapiens BAC clone RP11-725C19
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 64196)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 64196)
Waterston, R.H.
Direct Submission
JOURNAL Submitted (06-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 64196)
Waterston, R.H.
Direct Submission
JOURNAL Submitted (21-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 64196)
Waterston, R.
Direct Submission
JOURNAL Submitted (30-MAR-2002) Department of Genetics, Washington

COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 21, 2002 this sequence version replaced gi:19073866.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0725C19

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-8p17, 2000 bp overlap.
Actual start of this clone is at base position 185268 of RP11-393A12; actual end is at base position 64196 of RP11-725C19.

Polymorphisms exist between AC013365 and AC110804.

FEATURES
source

Location/Qualifiers
1. .64196
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-725C19"
/clone_lib="RPCI-11"
/rpt_family="AT_rich"
162. .182
/rpt_family="AT_rich"
312. .607
/rpt_family="Alu"
369. .505
/note="match to EST BG257959 (NID:gl2767775)"
877. .909
/rpt_family="AT_rich"
1113. .1536
/rpt_family="L2"
2400. .2427
/rpt_family="AT_rich"
2941. .2984
/rpt_family="L2"
3413. .3435
/rpt_family="(TA)n"
3829. .4216

repeat_region
repeat_region
repeat_region
misc_feature
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
misc_feature

misc_feature /note="match to EST H06135 (NID:g869687) y177b12.s1"
3829. .4170
misc_feature /note="match to EST BI497113 (NID:g15336457)"
3831. .4157
misc_feature /note="match to EST AA639698 (NID:g2563477) np60f09.s1"
3835. .4360
misc_feature /note="match to EST AW237145 (NID:g6569534) xm66b03.x1"
3836. .4376
misc_feature /note="similar to EST BM791320 (NID:g19139552)"
3837. .4405
misc_feature /note="similar to Bos taurus EST BI539797 (NID:g15380907)"
3844. .4610
misc_feature /note="match to EST BI857800 (NID:g15998547)"
3933. .4394
misc_feature /note="match to EST AI694383 (NID:g4971723) wd83b04.x1"
4105. .4843
misc_feature /note="match to EST BG177619 (NID:g12684322)"
4137. .4527
misc_feature /note="match to EST AI267688 (NID:g3886855) aq93c10.x1"
4153. .4433
misc_feature /note="match to EST C00538 (NID:g1432768)"
4233
misc_feature /note="match to EST BI497113 (NID:g15336457)"
4359. .4746
misc_feature /note="match to EST AA227760 (NID:g1849333) zr56a07.r1"
4404. .5178
misc_feature /note="match to EST BG291471 (NID:g13049386)"
4412. .4619
misc_feature /note="similar to Homo sapiens EST BF154977
(NID:g11050160)"
4657. .5028
misc_feature /note="match to EST R20307 (NID:g774941) yg35a09.r1"
4727. .5195
misc_feature /note="similar to Bos taurus EST AW669076 (NID:g7525590)"
4740. .5803
misc_feature /note="similar to EST BM547723 (NID:g18781736)"
4859. .5265
misc_feature /note="match to EST AW901105 (NID:g8065310)"
5002. .5687
misc_feature /note="match to EST BG254117 (NID:g12763933)"
5004. .5213
misc_feature /note="match to EST BF084219 (NID:g10878049)"
5074. .5447
misc_feature /note="match to EST N80420 (NID:g1243121) yz92e06.r1"
5255. .6002
misc_feature /note="match to EST BI869486 (NID:g16043159)"
5442. .6151
misc_feature /note="match to EST AW961203 (NID:g8150887)"
5497. .5830
misc_feature /note="match to EST AI540897 (NID:g4458270)"
5782. .6124
misc_feature /note="match to EST AA309186 (NID:g1961511)"
5875. .6701
misc_feature /note="match to EST BI334972 (NID:g15019629)"
5995. .6331
misc_feature /note="match to EST AA471165 (NID:g2199022)"
6000. .6124
misc_feature /note="match to EST AW996896 (NID:g8257130)"
6032. .6836
misc_feature /note="similar to EST BM474943 (NID:g18523985)"
6140. .6518
misc_feature /note="match to EST AW996897 (NID:g8257131)"
6444. .6885
misc_feature /note="match to EST AA375957 (NID:g2028275)"
6878. .7187
repeat_region /rpt_family="Alu"
6899. .7362
misc_feature /note="match to EST AA488687 (NID:g2216118) ab38f03.s1"
6970. .7746
misc_feature /note="match to EST BE739933 (NID:g10153837)"
7099. .7815
misc_feature /note="match to EST BF381953 (NID:g11363256)"
7124. .7125

misc_feature /note="match to EST AW996897 (NID:g8257131)"
7174. .7495
misc_feature /note="match to EST BF927528 (NID:g12325656)"
7253. .7884
misc_feature /note="similar to Homo sapiens EST BG499179
(NID:g13460696)"
7568. .7949
misc_feature /note="match to EST AA307001 (NID:g1959330)"
7613. .8332
misc_feature /note="similar to EST BM706727 (NID:g19019985)"
7836. .8011
repeat_region /rpt_family="Achobo"
8035. .8572
misc_feature

Query Match 21.4%; Score 479; DB 9; Length 64196;
Best Local Similarity 99.8%; Pred. No. 4.5e-261;
Matches 529; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1671 AGAGAAAATAACGAGAACATTACAAATAATACTGGAAGTTGTACCAGAGAAGATAAGTT 1730
|||||
Db 11755 AGAGAAAATAACGAGAACATTACAAATAATACTGGAAGTTGTACCAGAGAAGATAAGTT 11696
|||||
QY 1731 ATGAACATAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACACAAAATAGGGATT 1790
|||||
Db 11695 ATGAACATAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACACAAAATAGGGATT 11636
|||||
QY 1791 TTTACTTCAATTTCTGAAAGTCTAGAGAAATTACAACCTTTGGTGATAAACAAAGGAGTCA 1850
|||||
Db 11635 TTTACTTCAATTTCTGAAAGTCTAGAGAAATTACAACCTTTGGTGATAAACAAAGGAGTCA 11576
|||||
QY 1851 GTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCTAAGAAATTTAGTTATAA 1910
|||||
Db 11575 GTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCTAAGAAATTTAGTTATAA 11516
|||||
QY 1911 CTCTATGTAGTTATAGAAAAGTGAATATGCAGTTATCTATGAGTCGCACAATTTCTTGAGT 1970
|||||
Db 11515 CTCTATGTAGTTATAGAAAAGTGAATATGCAGTTATCTATGAGTCGCACAATTTCTTGAGT 11456
|||||
QY 1971 CTCTGATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAAATTTACTATGTGGTCATTC 2030
|||||
Db 11455 CTCTGATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAAATTTACTATGTGGTCATTC 11396
|||||
QY 2031 TCTACAACATATGTTAGCACGGCAAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTA 2090
|||||
Db 11395 TCTACAACATATGTTAGCACGGCAAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTA 11336
|||||
QY 2091 TATATGGGTTTTGTTAAAGATGGTTTTTACACACTACAGATGCTCTATACTGTGAAAAGTGT 2150
|||||
Db 11335 TATATGGGTTTTGTTAAAGATGGTTTTTACACACTATAGATGCTCTATACTGTGAAAAGTGT 11276
|||||
QY 2151 TTCAATTTCTGAAAAAAGCATACATCATGATGATGTCGCAAGAGGAGAGAA 2200
|||||
Db 11275 TTCAATTTCTGAAAAAAGCATACATCATGATGATGTCGCAAGAGGAGAGAA 11226
|||||

RESULT 10
AX321911
LOCUS AX321911 337 bp DNA linear PAT 15-DEC-2001
DEFINITION Sequence 442 from Patent WO0172295.
ACCESSION AX321911
VERSION AX321911.1 GI:17906521
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Reed,S.G.; Lodes,M.J.; Mohamath,R.; Secrist,H.; Benson,D.R.;
Indirias,C.Y.; Henderson,R.A.; Fling,S.P.; Algate,P.A.; Elliot,M.,
Mannion,J. and Kalos,M.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0172295-A 442 04-OCT-2001;

FEATURES
source
CORIXA CORPORATION (US)
Location/Qualifiers
1. .337
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 103 a 60 c 93 g 81 t
ORIGIN

Query Match 12.6%; Score 281; DB 6; Length 337;
Best Local Similarity 100.0%; Pred. No. 4.1e-148;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAGGTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAC 60
|||||
Db 4 GGAGGTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAC 63
|||||

QY 61 AGAATTAGGAAAAAGAAAGAAAAAGAGAGAGAGAGAAATCCAGGCCAATTGTGG 120
|||||
Db 64 AGAATTAGGAAAAAGAAAGAAAAAGAGAGAGAGAGAAATCCAGGCCAATTGTGG 123
|||||

QY 121 CATAGATTTATCATATTTCTGGATTTTGGATTTCTTTTCTCATCATCGGATTC 180
|||||
Db 124 CATAGATTTATCATATTTCTGGATTTTGGATTTCTTTTCTCATCATCGGATTC 183
|||||

QY 181 GGAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGG 240
|||||
Db 184 GGAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGG 243
|||||

QY 241 AGGCTGCTTCCCTGGCAACAAGGAGCCACCTGGGCAGGA 281
|||||
Db 244 AGGCTGCTTCCCTGGCAACAAGGAGCCACCTGGGCAGGA 284
|||||

RESULT 11
AX351341/c
LOCUS
DEFINITION Sequence 88 from Patent WO0196390.
ACCESSION AX351341
VERSION AX351341.1 GI:18616688
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196390-A 88 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1. .596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 160 a 151 c 126 g 158 t 1 others
ORIGIN

Query Match 12.0%; Score 269; DB 6; Length 596;
Best Local Similarity 100.0%; Pred. No. 2.9e-141;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
|||||
Db 269 GAGAAAGTCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTTGGC 210
|||||

QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGCTCCAGAACACGGGCAGC 453
|||||
Db 209 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGCTCCAGAACACGGGCAGC 150
|||||

QY 454 GTGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTTGGAGCTTTG 513
|||||
Db 149 GTGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTTGGAGCTTTG 90
|||||

QY 514 TCTTATGCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCAATACACATATATTTG 573
|||||
Db 89 TCTTATGCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCAATACACATATATTTG 30
|||||

QY 574 GAAGTCTTTGGTCCATTACCAAGCTTTTGT 602
|||||
Db 29 GAAGTCTTTGGTCCATTACCAAGCTTTTGT 1
|||||

RESULT 12
AC093903/c
LOCUS
DEFINITION Homo sapiens BAC clone RP11-733C7 from 4, complete sequence.
ACCESSION AC093903 AC055827
VERSION AC093903.3 GI:15920156
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 161280)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 161280)
AUTHORS Radionenko,M. and Kozlowicz,A.
TITLE The sequence of Homo sapiens BAC clone RP11-733C7
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 161280)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 161280)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 161280)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Oct 4, 2001 this sequence version replaced gi:15625016.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0733C07
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.

QY 61 AGAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 120
|||||
Db 27696 AGAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 27637
|||||
QY 121 CATAGATTTTATCATATCTTGGAATTTTGGATCTTTTGTCTCATCACTGGATTCA 180
|||||
Db 27636 CATAGATTTTATCATATCTTGGAATTTTGGATCTTTTGTCTCATCACTGGATTCA 27577
|||||
QY 181 GG 182
||
Db 27576 GG 27575

RESULT 13
LOCUS AB042201 1683 bp DNA linear PRI 10-FEB-2001
DEFINITION Homo sapiens xCT gene for cystine/glutamate transporter, exon 1,
partial cds.
ACCESSION AB042201
VERSION AB042201.1 GI:7670267
KEYWORDS cystine/glutamate transporter.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Sato,H., Tamba,M., Kuriyama-Matsumura,K., Okuno,S. and Bannai,S.
TITLE Molecular cloning and expression of human xCT, the light chain of
amino acid transport system xc-
JOURNAL Antioxidants and Redox Signaling 2, 665-671 (2000)
REFERENCE 2 (bases 1 to 1683)
AUTHORS Sato,H. and Bannai,S.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Hideyo Sato, University of Tsukuba,
Institute of Basic Medical Sciences; Tennodai 1-1-1, Tsukuba,
Ibaraki 305-8575, Japan (E-mail:hideyo-semd.tsukuba.ac.jp,
Tel:81-298-53-3282, Fax:81-298-53-3039)
FEATURES
source Location/Qualifiers
1..1683
/organism="Homo sapiens"
/db_xref="taxon:9606"
1001..1683
/gene="xCT"
1001..1555
/gene="xCT"
/product="cystine/glutamate transporter"
/number=1
1279..>1555
/gene="xCT"
/codon_start=1
/product="cystine/glutamate transporter"
/protein_id="BAA94999.1"
/db_xref="GI:7670268"
/translation="MVRKPVVSTISKGYLQGNVNGRLPSLGNKEPPGQEKVQLKRKV
TLRGVSIIGTIIIGAGIFISPKGVQLQNTGSVGNSLTIWTVCGVLSLF"
1556..>1683
/gene="xCT"
/number=1
BASE COUNT 495 a 287 c 370 g 531 t
ORIGIN
Query Match 7.7%; Score 173; DB 9; Length 1683;
Best Local Similarity 100.0%; Pred. No. 1.9e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
|||||
Db 1384 GAGAAAGTGCAGTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 1443
|||||
QY 394 ACCATCATTTGGACGAGGAATCTTTCATCTCTCCTTAAGGGCGTCTCCAGAACACGGGCAGC 453
|||||
Db 1444 ACCATCATTTGGACGAGGAATCTTTCATCTCTCCTTAAGGGCGTCTCCAGAACACGGGCAGC 1503
|||||

QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCTGTCTCACTATTGG 506
|||||
Db 1504 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCTGTCTCACTATTGG 1556
|||||
RESULT 14
LOCUS AC116610/c 3994 bp DNA linear PRI 29-MAY-2002
DEFINITION Homo sapiens BAC clone RP13-562H2 from 4, complete sequence.
ACCESSION AC116610
VERSION AC116610.4 GI:20279502
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3994)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 3994)
AUTHORS Desai,A. and Cotton,M.
TITLE The sequence of Homo sapiens BAC clone RP13-562H2
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 3994)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 3994)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (13-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 3994)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 6 (bases 1 to 3994)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 24, 2002 this sequence version replaced gi:20136935.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_FH0562H02

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-13 Human Female BAC Library was constructed using improved cloning techniques developed by Kazutoyo Osoegawa. The library was generated by Baohui Zhao in our laboratory. Construction was funded by a grant from the National Human Genome Research Institute (NHGRI, NIH) (#1R01HG01165-03). The library was generated according to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA Sequencing'.

Female blood was obtained via a double-blind selection protocol. Female blood DNA was isolated from one randomly chosen donor (out of 10 female donors) and partially digested with a combination of EcoRI and EcoRI Methylase for library segments 1&2 or either MboI or DpnII for library segments 3&4. Size selected DNA was cloned into the pBACE3.6 vector between the EcoRI sites for library segments 1&2 or the BamHI sites for library segments 3&4. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). The library has been arrayed into 384-well microtiter dishes and also gridded onto 22x22cm nylon high density filters for screening by probe hybridization.

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-725C19, 2000 bp overlap; the clone sequenced to the right is RP11-733C7, 2000 bp overlap. Actual start of this clone is at base position 92698 of RP11-8P17; actual end is at base position 68420 of RP11-733C7.

A single plasmid subclone region exists between bases 2382-2397. Polymorphisms have been identified between AC110804, AC093903, AC105310, and AC116610.

FEATURES

source
1. .3994
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP13-562H2"
/clone_lib="RPCI-13"
169. .205
/rpt_family="AT_rich"
935. .1089
/rpt_family="MER1_type"
1606. .1716
/rpt_family="L2"
2142. .2298
/rpt_family="L1"
2300. .2400
/rpt_family="(GGAA)n"
2465. .2519
/rpt_family="AT_rich"
2524. .2678
/rpt_family="L1"
2679. .2716
/rpt_family="(TA)n"
2717. .2786
/rpt_family="L1"
3387. .3476
/rpt_family="GA-rich"
3701. .3840
/rpt_family="L1"
3841. .3960
/rpt_family="L1"
BASE COUNT 1431 a 656 c 689 g 1218 t
ORIGIN

Query Match 5.7%; Score 128; DB 9; Length 3994;
Best Local Similarity 100.0%; Pred. No. 9.5e-61;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 749 CTGTAGTGATGGTCCTAAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTTCT 808
|||||
Db 864 CTGTAGTGATGGTCCTAAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTTCT 805
|||||
QY 809 TAACCTTTTGTCAAGCTCACAGCAATCTGTATAATTATAGTCCCTGGAGTTATGAGCTAA 868
|||||
Db 804 TAACCTTTTGTCAAGCTCACAGCAATCTGTATAATTATAGTCCCTGGAGTTATGAGCTAA 745
|||||
QY 869 TTAAGGT 876
|||||
Db 744 TTAAGGT 737
|||||
RESULT 15
AX341060 575 bp DNA linear PAT 10-JAN-2002
LOCUS AX341060
DEFINITION Sequence 1307 from Patent WO0196388.
ACCESSION AX341060
VERSION AX341060.1 GI:18137042
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang, Y., Harlocker, S.L. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196388-A 1307 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source 1. .575
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 151 a 122 c 153 g 141 t 8 others
ORIGIN
Query Match 4.8%; Score 107; DB 6; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.8e-49;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
|||||
Db 354 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC 413
|||||
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCCA 440
|||||
Db 414 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCCA 460
|||||

Search completed: April 16, 2003, 13:16:27
Job time : 6862 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 10:51:38 ; Search time 3547 Seconds
(without alignments)
10223.201 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	455	20.3	516	9 AL702210	AL702210 DKFZp686C
C 2	450	20.1	483	10 AW205572	AW205572 UI-H-B11-
C 3	438	19.6	511	14 BM887840	BM887840 TMT023 Hu
4	429	19.2	919	12 BG490131	BG490131 602519050
5	262	910	11.7	12 BG326527	BG326527 602425373
6	236	10.5	936	12 BG284503	BG284503 602408645

C 7	172	7.7	557	17	AQ002318	AQ002318 CIT-HSP-2
8	171	7.6	371	14	W00842	W00842 yz96d12.r1
9	153	6.8	400	17	AQ533681	AQ533681 RPCI-11-3
10	153	6.8	486	17	B33331	B33331 HS-1017-A1-
11	152	6.8	497	14	BQ370588	BQ370588 PMO-GN034
12	128	5.7	485	17	BH860963	BH860963 UP_336-3L
13	112	5.0	1072	12	BG388107	BG388107 602413070
C 14	102	4.6	752	12	BG186884	BG186884 RST5861 A
C 15	87	3.9	87	10	AW059922	AW059922 LFIb04.Yg
16	84	3.8	632	17	AQ201392	AQ201392 RPCI11-60
17	78	3.5	388	9	AA812721	AA812721 aj31g09.s
C 18	52	2.3	530	9	AI102370	AI102370 EST211659
C 19	51	2.3	123	9	AI313891	AI313891 hpi-7 PMA
20	50	2.2	869	10	BE535700	BE535700 601060361
C 21	49	2.2	235	14	BQ448503	BQ448503 UI-H-EU1-
C 22	49	2.2	255	10	AW449544	AW449544 UI-H-BI3-
C 23	49	2.2	417	13	BM662461	BM662461 UI-E-CI1-
C 24	49	2.2	602	13	BM511762	BM511762 1j53b05.x
C 25	49	2.2	627	10	AV704488	AV704488 AV704488
C 26	49	2.2	668	14	BM973417	BM973417 UI-CF-EC1
27	49	2.2	688	10	AV706925	AV706925 AV706925
28	49	2.2	1028	14	BQ428200	BQ428200 AGENCOURT
29	45	2.0	179	12	BF031033	BF031033 601558910
C 30	45	2.0	237	12	BE857736	BE857736 7g47e06.x
C 31	45	2.0	262	14	N72154	N72154 yz99a03.s1
32	45	2.0	270	14	F15607	F15607 HSPD000073 H
C 33	45	2.0	286	9	AI125945	AI125945 qc41h09.x
C 34	45	2.0	286	9	AI375209	AI375209 tc10f02.x
C 35	45	2.0	324	9	AA722121	AA722121 zh17h07.s
C 36	45	2.0	343	9	AA922752	AA922752 om50h04.s
C 37	45	2.0	343	9	AI123247	AI123247 qa47f09.x
C 38	45	2.0	350	12	BF095483	BF095483 IL2-UT007
C 39	45	2.0	358	9	AA398455	AA398455 zt47g08.s
40	45	2.0	359	14	F37452	F37452 HSPD36219 H
C 41	45	2.0	364	9	AI424212	AI424212 tf63e10.x
C 42	45	2.0	372	14	BQ013223	BQ013223 UI-1-BC1p
C 43	45	2.0	402	14	BQ017470	BQ017470 UI-H-DT1-
C 44	45	2.0	409	9	AA992074	AA992074 ot37c10.s
C 45	45	2.0	422	9	AI753931	AI753931 cr16c05.x

ALIGNMENTS

RESULT 1
AL702210
LOCUS AL702210 516 bp mRNA linear EST 22-MAR-2002
DEFINITION DKFZp686C09156_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DKFZp686C09156 5', mRNA sequence.
ACCESSION AL702210 GI:196855565
VERSION AL702210.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 516)
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: Bloecker H
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686C09156) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. 516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686C09156"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"

BASE COUNT 107 a 121 c 110 g 178 t
ORIGIN

Query Match 20.3%; Score 455; DB 9; Length 516;
Best Local Similarity 100.0%; Pred. No. 6.1e-125;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 AAAACCCCTGAAAAAACCAATCCCCCTTGCAATATGATATCCATGGCCATTTGTCACCATTTG 1066
|||||
Db 62 AAAACCCCTGAAAAAACCAATCCCCCTTGCAATATGATATCCATGGCCATTTGTCACCATTTG 121
QY 1067 GCTATGTGCTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTT 1126
|||||
Db 122 GCTATGTGCTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTT 181
QY 1127 CAAATGCGAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCAATAGCAGTTC 1186
|||||
Db 182 CAAATGCGAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCAATAGCAGTTC 241
QY 1187 CGATCTTTGTTGCCCTCTCCCTCTGCTCCATGACGAGTGGTGTGTTGCTGTCTCCA 1246
|||||
Db 242 CGATCTTTGTTGCCCTCTCCCTCTGCTCCATGACGAGTGGTGTGTTGCTGTCTCCA 301
QY 1247 GGTATTCTATGTTGCGTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATG 1306
|||||
Db 302 GGTATTCTATGTTGCGTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATG 361
QY 1307 TCCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTTCACCCCTTTGACAAATGATAATGC 1366
|||||
Db 362 TCCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTTCACCCCTTTGACAAATGATAATGC 421
QY 1367 TCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTA 1426
|||||
Db 422 TCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTA 481
QY 1427 TTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGA 1461
|||||
Db 482 TTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGA 516

RESULT 2
AW205572/c
LOCUS
DEFINITION UI-H-BI1-adt-h-02-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
IMAGE:2718171 3', mRNA sequence.

ACCESSION AW205572
VERSION AW205572.1 GI:6505046
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 483)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1. 483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2718171"
/clone_lib="NCI_CGAP_Sub3"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NCI_CGAP_Sub3 library is a subtracted library derived from the NCI_CGAP_Sub1 library, which is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Lei2, NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_LIB=NCI_CGAP_Co10
TAG_TISSUE=colon
TAG_SEQ=AAACG"

BASE COUNT 150 a 98 c 98 g 137 t
ORIGIN

Query Match 20.1%; Score 450; DB 10; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.9e-123;
Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1394 TGAATTTCTCAGTTTGGCCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTT 1453
|||||
Db 459 TGAATTTCTCAGTTTGGCCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTT 400
QY 1454 ATCTTCGATACAAATGCCAGATATGCATCGTCCTTTCAAGGTGCCACTGTTTCATCCAG 1513
|||||
Db 399 ATCTTCGATACAAATGCCAGATATGCATCGTCCTTTCAAGGTGCCACTGTTTCATCCAG 340
QY 1514 CTTTGTTCCTTCACATGCCCTTCATGGTTGCCCTTTCCCTCTATTTCGGACCCATTTA 1573
|||||
Db 339 CTTTGTTCCTTCACATGCCCTTCATGGTTGCCCTTTCCCTCTATTTCGGACCCATTTA 280
QY 1574 GTACAGGGATGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATTA 1633
|||||
Db 279 GTACAGGGATGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATTA 220

```
QY 1634 TATGGGACAGAAACCCAGGTGGTTAGAAATAATGTCAGAGAAAATAACCCAGAACATTAC 1693
Db 219 TATGGGACAGAAACCCAGGTGGTTAGAAATAATGTCAGAGAAAATAACCCAGAACATTAC 160

QY 1694 AAATAATACTGGGAAGTTGTACCCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCT 1753
Db 159 AAATAATACTGGGAAGTTGTACCCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCT 100

QY 1754 TGGCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTTCATTTTCTGAAAGTCT 1813
Db 99 TGGCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTTCATTTTCTGAAAGTCT 40

QY 1814 AGAGAAATTACAACCTTTGGTGATAAACAAA 1843
Db 39- AGAGAAATTACAACCTTTGGTGATAAACAAA 10

RESULT 3
BM887840/c
LOCUS BM887840 511 bp mRNA linear EST 08-MAR-2002
DEFINITION TMT023 Human Trabecular Meshwork cDNA library Homo sapiens cDNA 5',
mRNA sequence.
ACCESSION BM887840
VERSION BM887840.1 GI:19271584
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 511)
AUTHORS Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
TITLE Expression Profile and Genome Location of cDNA Clones from an
JOURNAL Infant Human Trabecular Meshwork Library
COMMENT Unpublished (2002)
Contact: Wirtz MK
Glaucoma Genetics Lab
Oregon Health Sciences University
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
Tel: 503-494-4698
Fax: 503-494-6875
Email: wirtzm@ohsu.edu
Seq primer: T7 Reverse.
Location/Qualifiers
1. 511
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human Trabecular Meshwork cDNA library"
/tissue_type="eye"
/cell_type="trabecular meshwork"
/dev_stage="2 week to 2 year old infants"
/lab_host="TOP10F"
/note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
cDNA library made from mRNA isolated from trabecular
meshwork cells established from eyes from 6 individuals,
ages 2 weeks to 2 years. Cells were harvested at passages
3 through 6. Invitrogen made a unidirectional cDNA library
from the mRNA from the frozen cells using a pCDNA3 vector
and TP010F", host cells."
BASE COUNT 161 a 99 c 70 g 180 t 1 others
ORIGIN

Query Match 19.6%; Score 438; DB 14; Length 511;
Best Local Similarity 99.8%; Pred. No. 6.5e-120;
Matches 488; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1636 TGGGACAGAAACCCAGGTGGTTAGAAATAATGTCAGAGAAAATAACCCAGAACATTACAA 1695
Db 511 TGGGACAGAAACCCAGGTGGTTAGAAATAATGTCAGAGAAAATAACCCAGAACATTACAA 452

QY 1696 ATAATACTGGGAAGTTGTACCCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCTTG 1755
Db 451 ATAATACTGGGAAGTTGTACCCAGAAGAAGATAAGTTATGAACATAATGGACTTGAGATCTTG 392
```

```
QY 1756 GCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCTCAATTTCTGAAAGTCTAG 1815
Db 391 GCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCTCAATTTCTGAAAGTCTAG 332

QY 1816 AGAATTACAACCTTTGGTGATAAAACAAAAGGAGCTCAGTTATTTTATTCATATATTTTAGC 1875
Db 331 AGAATTACAACCTTTGGTGATAAAACAAAAGGAGCTCAGTTATTTTATTCATATATTTTAGC 272

QY 1876 ATATTCGAACATAATTCTAAGAAATTTAGTTATAACTCTATGTAGTATAGAAAGTGAAT 1935
Db 271 ATATTCGAACATAATTCTAAGAAATTTAGTTATAACTCTATGTAGTATAGAAAGTGAAT 212

QY 1936 ATGCAGTTATTCTATGAGTCGCCACAATTTCTGAGTCTCTGATACCTACCTATTTGGGGTTA 1995
Db 211 ATGCAGTTATTCTATGAGTCGCCACAATTTCTGAGTCTCTGATACCTACCTATTTGGGGTTA 152

QY 1996 GGAGAAAAGACTAGACAATTACTATGTGGTCATTTCTTACAACATATGTTAGCACGGCAA 2055
Db 151 GGAGAAAAGACTAGNCAATTACTATGTGGTCATTTCTTACAACATATGTTAGCACGGCAA 92

QY 2056 AGAACCTTCAAATGAAGACTGAGATTTTCTGTATATATATGTTTGTAAAGATGGTTT 2115
Db 91 AGAACCTTCAAATGAAGACTGAGATTTTCTGTATATATATGTTTGTAAAGATGGTTT 32

QY 2116 TACACACTA 2124
Db 31 TACACACTA 23

RESULT 4
BG490131
LOCUS BG490131 919 bp mRNA linear EST 27-MAR-2001
DEFINITION 602519050F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4637509 5',
mRNA sequence.
ACCESSION BG490131
VERSION BG490131.1 GI:13451641
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 919)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/Drp/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI398 row: c column: 14
High quality sequence stop: 726.
Location/Qualifiers
1. 919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4637509"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

```
BASE COUNT      253 a      195 c      199 g      272 t
ORIGIN

Query Match      19.2%; Score 429; DB 12; Length 919;
Best Local Similarity 100.0%; Pred. No. 1.8e-117;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1344 GCACCCCTTGACAATGATAATGCTCTCTCTGGAGACCTGCACAGTCTTTTGAATTCCT 1403
Db      |||||||
QY 297  GCACCCCTTGACAATGATAATGCTCTCTCTGGAGACCTGCACAGTCTTTTGAATTCCT 356
Db      |||||||

QY 1404 CAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATA 1463
Db      |||||||

QY 357  CAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATA 416
Db      |||||||

QY 1464 CAAATGCCCAGATATGCATCGTCCCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTC 1523
Db      |||||||

QY 417  CAAATGCCCAGATATGCATCGTCCCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTC 476
Db      |||||||

QY 1524 CTTACATGCCTCTTCATGGTTGCCCTTCCCTCTATTCCGACCCCATTTAGTACAGGAT 1583
Db      |||||||

QY 477  CTTACATGCCTCTTCATGGTTGCCCTTCCCTCTATTCCGACCCCATTTAGTACAGGAT 536
Db      |||||||

QY 1584 TGGCTTCGTCATCACTCTGACTGGAGTCCCTCGGTATTATCTCTTATTATATGGGACAA 1643
Db      |||||||

QY 537  TGGCTTCGTCATCACTCTGACTGGAGTCCCTCGGTATTATCTCTTATTATATGGGACAA 596
Db      |||||||

QY 1644 GAAACCCAGGTGGTTTAGAATAATGTCAGAGAGAAATAACCCAGAACATTACAAATAACT 1703
Db      |||||||

QY 597  GAAACCCAGGTGGTTTAGAATAATGTCAGAGAGAAATAACCCAGAACATTACAAATAACT 656
Db      |||||||

QY 1704 GGAAGTTGTACCAAGAAGAAGATAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTG 1763
Db      |||||||

QY 657  GGAAGTTGTACCAAGAAGAAGATAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTG 716
Db      |||||||

QY 1764 CCCAAGGGG 1772
Db      |||||||

QY 717  CCCAAGGGG 725
Db      |||||||

RESULT 5
BG326527      910 bp      mRNA      linear      EST 27-FEB-2001
LOCUS      602425373F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562994 5',
DEFINITION      mRNA sequence.
ACCESSION      BG326527
VERSION      BG326527.1 GI:13132964
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 910)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1276 row: b column: 19
High quality sequence stop: 706.
Location/Qualifiers
1. .910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4562994"
/clone_lib="NIH_MGC_14"
```

```
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II-RT (Life Technologies)."
```

BASE COUNT 237 a 178 c 249 g 246 t
ORIGIN

Query Match 11.7%; Score 262; DB 12; Length 910;
Best Local Similarity 99.7%; Pred. No. 3.8e-68;
Matches 382; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
Db |||||||
QY 265 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGC 324
Db |||||||

QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 453
Db |||||||

QY 325 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTCTCCAGAACACGGGCAGC 384
Db |||||||

QY 454 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGCTCACTATTTGGAGCTTTG 513
Db |||||||

QY 385 GTGGGCATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGCTCACTATTTGGAGCTTTG 444
Db |||||||

QY 514 TCTTATGCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCACTACACATATATTTTG 573
Db |||||||

QY 445 TCTTATGCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCACTACACATATA-TTTG 503
Db |||||||

QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATAATACGC 633
Db |||||||

QY 504 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATAATACGC 563
Db |||||||

QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 693
Db |||||||

QY 564 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 623
Db |||||||

QY 694 ATTCAATGTGAAATCCCTGAACT 716
Db |||||||

QY 624 ATTCAATGTGAAATCCCTGAACT 646
Db |||||||

RESULT 6
BG284503 936 bp mRNA linear EST 21-FEB-2001
LOCUS 602408645F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4537810 5',
DEFINITION mRNA sequence.
ACCESSION BG284503
VERSION BG284503.1 GI:13035516
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 936)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM10463 row: i column: 11
High quality sequence stop: 795.

FEATURES
source

Location/Qualifiers
1. .936
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4537810"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 248 a 202 c 251 g 235 t
ORIGIN

Query Match 10.5%; Score 236; DB 12; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.8e-60;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
|||||
Db 376 GAGAAAGTCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTTGGC 435
|||||
QY 394 ACCATCATTTGAGCAGGAAATCTTCATCTCTCCTAAGGGCGTGCTCCAGAACACGGGCAGC 453
|||||
Db 436 ACCATCATTTGAGCAGGAAATCTTCATCTCTCCTAAGGGCGTGCTCCAGAACACGGGCAGC 495
|||||
QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTGGAGCTTTG 513
|||||
Db 496 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTGGAGCTTTG 555
|||||

QY 514 TCTTATGCTGAATTGGGAACAACACTATAAGAAATCTGGAGGTCAATTACACATATAT 569
|||||
Db 556 TCTTATGCTGAATTGGGAACAACACTATAAGAAATCTGGAGGTCAATTACACATATAT 611
|||||

RESULT 7

AQ002318/c
LOCUS
DEFINITION
CIT-HSP-2283E8.TF CIT-HSP Homo sapiens genomic clone 2283E8, DNA
sequence.
ACCESSION
AQ002318
VERSION
AQ002318.1 GI:3029522
KEYWORDS
GSS.
SOURCE
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 557)

REFERENCE

AUTHORS
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.

TITLE

Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)

JOURNAL

COMMENT
Unpublished (1998)
Other_GSSs: CIT-HSP-2283E8.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. .557
/organism="Homo sapiens"
/db_xref="GDB:7147907"

/db_xref="taxon:9606"
/clone="2283E8"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 142 a 152 c 133 g 130 t
ORIGIN

Query Match 7.7%; Score 172; DB 17; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.2e-41;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 AGAAAGTCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTTGGCA 394
|||||
Db 278 AGAAAGTCAGCTGAAGAGGAAAGTCACCTTTACTGAGGGGAGTCTCCATTATCATTTGGCA 219
|||||
QY 395 CCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTGCTCCAGAACACGGGCAGCG 454
|||||
Db 218 CCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTGCTCCAGAACACGGGCAGCG 159
|||||
QY 455 TGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTGG 506
|||||
Db 158 TGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTGG 107
|||||

RESULT 8

W00842
LOCUS

DEFINITION

Yz96dl2.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
IMAGE:290903 5', mRNA sequence.

ACCESSION

W00842

VERSION

W00842.1 GI:1272822

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 371)

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.

The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilton RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: ETprimer

High quality sequence stop: 339.

Location/Qualifiers

1. .371

/organism="Homo sapiens"

/db_xref="GDB:388403"

/db_xref="taxon:9606"

/clone="IMAGE:290903"

/clone_lib="Soares melanocyte 2NbHM"

/sex="Male"

/tissue_type="melanocyte"

/lab_host="DH10B (ampicillin resistant)"

/note="Vector: pT7T3D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGAGCGGCGAGTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector

FEATURES

source

```
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT      131 a      50 c      69 g      115 t      6 others
ORIGIN

Query Match      7.6%; Score 171; DB 14; Length 371;
Best Local Similarity 99.1%; Pred. No. 6.2e-41;
Matches 321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1661 GAATAATGTCAGAGAAAATAACCAACATTACAAATAATACTGGAAGTTGTACCAGAAG 1720
Db      |||||||
1 GAATAATGTCAGAGAAAATAACCAACATTACAAATAATACTGGAAGTTGTACCAGAAG 60

QY 1721 AAGATAAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAA 1780
Db      |||||||
61 AAGATAAGTTATGAACCTAATGGNCTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAA 120

QY 1781 AATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTGATAAACA 1840
Db      |||||||
121 AATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTGATAAACA 180

QY 1841 AAAGGAGTCAGTATTTTATTTCATATATTTTAGCATATTTTCAACTAATTTCTAAGAAAT 1900
Db      |||||||
181 AAAGGAGTCAGTATTTTATTTCATATATTTTAGCATATTTTCAACTAATTTCTAAGAAAT 240

QY 1901 TTAGTTATACTCTATGTAGTTATAGAAAGTGAATATGCAGTATTTCTATGAGTCGCACA 1960
Db      |||||||
241 TTAGTTATANCTCTATGTAGTTATAGAAAGTGAATATGCAGTATTTCTATGAGTCGCACA 300

QY 1961 ATTCTTGAGTCTCTGATACCTACC 1984
Db      |||||||
301 ATTCTTGAGTCTCTGATACCTACC 324

RESULT 9
AQ533681      400 bp      DNA      linear      GSS 18-MAY-1999
LOCUS      RPCI-11-384K17.TV RPCI-11 Homo sapiens genomic clone RPCI-11-384K17
DEFINITION      , DNA sequence.
ACCESSION      AQ533681
VERSION      AQ533681.1 GI:4845371
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 400)
AUTHORS      Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
TITLE      Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL      Unpublished (1997)
COMMENT      Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
FEATURES      Location/Qualifiers
source      1..400
/organism="Homo sapiens"
/db_xref="GDB:7647328"
/db_xref="taxon:9606"

(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT      131 a      50 c      69 g      115 t      6 others
ORIGIN

Query Match      7.6%; Score 171; DB 14; Length 371;
Best Local Similarity 99.1%; Pred. No. 6.2e-41;
Matches 321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1661 GAATAATGTCAGAGAAAATAACCAACATTACAAATAATACTGGAAGTTGTACCAGAAG 1720
Db      |||||||
1 GAATAATGTCAGAGAAAATAACCAACATTACAAATAATACTGGAAGTTGTACCAGAAG 60

QY 1721 AAGATAAGTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAA 1780
Db      |||||||
61 AAGATAAGTTATGAACCTAATGGNCTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAA 120

QY 1781 AATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTGATAAACA 1840
Db      |||||||
121 AATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAATTACAACTTTGGTGATAAACA 180

QY 1841 AAAGGAGTCAGTATTTTATTTCATATATTTTAGCATATTTTCAACTAATTTCTAAGAAAT 1900
Db      |||||||
181 AAAGGAGTCAGTATTTTATTTCATATATTTTAGCATATTTTCAACTAATTTCTAAGAAAT 240

QY 1901 TTAGTTATACTCTATGTAGTTATAGAAAGTGAATATGCAGTATTTCTATGAGTCGCACA 1960
Db      |||||||
241 TTAGTTATANCTCTATGTAGTTATAGAAAGTGAATATGCAGTATTTCTATGAGTCGCACA 300

QY 1961 ATTCTTGAGTCTCTGATACCTACC 1984
Db      |||||||
301 ATTCTTGAGTCTCTGATACCTACC 324

RESULT 9
AQ533681      400 bp      DNA      linear      GSS 18-MAY-1999
LOCUS      RPCI-11-384K17.TV RPCI-11 Homo sapiens genomic clone RPCI-11-384K17
DEFINITION      , DNA sequence.
ACCESSION      AQ533681
VERSION      AQ533681.1 GI:4845371
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 400)
AUTHORS      Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
TITLE      Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL      Unpublished (1997)
COMMENT      Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
FEATURES      Location/Qualifiers
source      1..400
/organism="Homo sapiens"
/db_xref="GDB:7647328"
/db_xref="taxon:9606"
```

```
/clone="RPCI-11-384K17"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT      103 a      84 c      72 g      140 t      1 others
ORIGIN

Query Match      6.8%; Score 153; DB 17; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.2e-35;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1344 GCACCCCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTCCT 1403
Db      |||||||
171 GCACCCCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTCCT 230

QY 1404 CAGTTTGGCCAGGTGGCTTTTATTATGGGTGGCAGTTGCTGGGTGATTTATCTTCGATA 1463
Db      |||||||
231 CAGTTTGGCCAGGTGGCTTTTATTATGGGTGGCAGTTGCTGGGTGATTTATCTTCGATA 290

QY 1464 CAAATGCCAGATATGTCATCGTCTTTCAAGGT 1496
Db      |||||||
291 CAAATGCCAGATATGTCATCGTCTTTCAAGGT 323

RESULT 10
B33331      486 bp      DNA      linear      GSS 17-OCT-1997
LOCUS      HS-1017-A1-G10-MF.abi CIT Human Genomic Sperm Library C Homo
DEFINITION      sapiens genomic clone Plate=CT 792 Col=19 Row=M, DNA sequence.
ACCESSION      B33331
VERSION      B33331.1 GI:2532700
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 486)
AUTHORS      Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
TITLE      Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
JOURNAL      Unpublished (1997)
COMMENT      Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackron@u.washington.edu
Sequence Tagged Connector
Plate: CT 792 row: M column: 19
Class: BAC ends
High quality sequence stop: 486.
FEATURES      Location/Qualifiers
source      1..486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=CT 792 Col=19 Row=M"
/clone_lib="CIT Human Genomic Sperm Library C"
/sex="M"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      112 a      96 c      109 g      166 t      3 others
ORIGIN

Query Match      6.8%; Score 153; DB 17; Length 486;
Best Local Similarity 100.0%; Pred. No. 1e-35;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1344 GCACCCCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTCCT 1403
|||||
```

Db 62 GCACCCCTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTCCT 121
QY 1404 CAGTTTGGCAGGTGGCTTTTATATTGGGCTGGCAGTTGGCTGATTTATCTTCGATA 1463
|||||
Db 122 CAGTTTGGCAGGTGGCTTTTATATTGGGCTGGCAGTTGGCTGATTTATCTTCGATA 181
QY 1464 CAAATGCCAGATATGCATCGTCCTTTCAAGGT 1496
|||||
Db 182 CAAATGCCAGATATGCATCGTCCTTTCAAGGT 214

RESULT 11
BQ370588
LOCUS BQ370588 497 bp mRNA linear EST 21-MAY-2002
DEFINITION PM0-GN0344-271200-001-f12 GN0344 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ370588
VERSION BQ370588.1 GI:21046102
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 497)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM0&t2=PM0-GN0344-
271200-001-f12&t3=2000-12-27&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 106

High quality sequence stop: 117.

Location/Qualifiers

1. .497

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="GN0344"

/dev_stage="Adult"

/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

BASE COUNT 103 a 120 c 113 g 161 t

ORIGIN

Query Match 6.8%; Score 152; DB 14; Length 497;

Best Local Similarity 100.0%; Pred. No. 2e-35;

Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1210 TTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGA 1269
|||||

Db 146 TTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATCTATGTTGCGTCTCGA 205
|||||

QY 1270 GAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACACTCCTCTACCA 1329

Db 206 GAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACACTCCTCTACCA 265
|||||
QY 1330 GCTGTTATTGTTTGCACCCTTTTGACAATGAT 1361
|||||
Db 266 GCTGTTATTGTTTGCACCCTTTTGACAATGAT 297

RESULT 12
BH860963
LOCUS BH860963 485 bp DNA linear GSS 18-JUL-2002
DEFINITION UP_336-3L_SP6 RPC111 Human Male BAC Library Homo sapiens genomic
clone 336-3L, DNA sequence.

ACCESSION BH860963

VERSION BH860963.1 GI:21905687

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 485)

AUTHORS Cheung,V.G., Dalrymple,H.L., Narasimhan,S., Watts,J., Schuler,G.,
Raap,A.K., Morley,M. and Bruzel,A.

TITLE A resource of mapped human bacterial artificial chromosome clones

JOURNAL Genome Res. 9 (10), 989-993 (1999)

MEDLINE 99455100

COMMENT

Contact: Narasimhan SL, Morley M, Burdick J, Cheung VG
Department of Pediatrics

University of Pennsylvania

3516 Civic Center Blvd, ARC 516, Philadelphia, PA 19104, USA

Tel: 215 590 2664

Fax: 215 590 3709

Email: sandya2@mail.med.upenn.edu

Plate: 336 row: L column: 3

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .485

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="336-3L"

/clone_lib="RPC111 Human Male BAC Library"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBACE3.6; RPC111 Human Male BAC Library"

BASE COUNT 132 a 97 c 93 g 163 t

ORIGIN

Query Match 5.7%; Score 128; DB 17; Length 485;

Best Local Similarity 100.0%; Pred. No. 2.5e-28;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 749 CTGTAGTGATGGTCTTAATAGCATGAGTGTGAGCTGGAGCGCCCGGATCCAGATTTCCT 808
|||||

Db 249 CTGTAGTGATGGTCTTAATAGCATGAGTGTGAGCTGGAGCGCCCGGATCCAGATTTCCT 308
|||||

QY 809 TAACCTTTGCAAGCTCACAGCAATTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAA 868
|||||

Db 309 TAACCTTTGCAAGCTCACAGCAATTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAA 368
|||||

QY 869 TTTAAAGGT 876

|||||

Db 369 TTTAAAGGT 376

RESULT 13

BG388107

LOCUS BG388107 1072 bp mRNA linear EST 12-MAR-2001

DEFINITION 602413070F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4521736 5',
mRNA sequence.

ACCESSION BG388107

VERSION BG388107.1 GI:13281553

KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1072)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10421 row: k column: 17
High quality sequence stop: 625.
Location/Qualifiers
1..1072
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4521736"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 331 a 229 c 270 g 242 t
ORIGIN
Query Match 5.0%; Score 112; DB 12; Length 1072;
Best Local Similarity 100.0%; Pred. No. 6.7e-24;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 391 GGCACCATCATGTGGAGCAGGAATCTTCATCTCTCTAAGGGGTCTCCAGAACACGGGC 450
Db 452 GGCACCATCATGTGGAGCAGGAATCTTCATCTCTCTAAGGGGTCTCCAGAACACGGGC 511
QY 451 AGCGTGGGCATCTCTTGACCATCTGGACGGTGTGTGGGGTCTGTCACTAT 502
Db 512 AGCGTGGGCATCTCTTGACCATCTGGACGGTGTGTGGGGTCTGTCACTAT 563
RESULT 14
BG186884/c
LOCUS BG186884 752 bp mRNA linear EST 21-APR-2001
DEFINITION RST5861 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG186884
VERSION BG186884.1 GI:13708571
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 752)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
High quality sequence stop: 441.
Location/Qualifiers
1..752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT 246 a 147 c 132 g 225 t
ORIGIN
Query Match 4.6%; Score 102; DB 12; Length 752;
Best Local Similarity 100.0%; Pred. No. 8.2e-21;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1983 CCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGGTCACTTCTACAACATAT 2042
Db 278 CCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGGTCACTTCTACAACATAT 219
QY 2043 GTTAGCACGGCAAGAACCTTCAATTTGAAGACTGAGATTTT 2084
Db 218 GTTAGCACGGCAAGAACCTTCAATTTGAAGACTGAGATTTT 177
RESULT 15
AW059922/c
LOCUS LF1b04.yg UPC15 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW059922
ACCESSION AW059922
VERSION AW059922.1 GI:6652244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 87)
AUTHORS Brenner,S., Williams,S.R., Vermass,E.H., Storck,T., Moon,K.,
McCollum,C., Mao,J.I., Kirchner,J.J., Eletr,S., DuBridge,R.B.,
Burcham,T. and Albrecht,G.
TITLE In vitro cloning of complex mixtures of DNA on microbeads: Physical
separation of differentially expressed cDNAs
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1665-1670 (2000)
MEDLINE 20144098
COMMENT Contact: Burcham TS
LYNX Therapeutics, Inc.
25861 Industrial Blvd., Hayward, CA 94545, USA
Tel: 510 670 9338
Fax: 510 670 9302
Email: timbelynxgen.com
Sequence obtained from LYNX Therapeutics Megasort technology.
Collected from the up-regulated gate.
High quality sequence stop: 87.
Location/Qualifiers
1..87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UPC15"
/cell_type="monocytic leukemia"
/note="Vector: pCR2.1; Cloning of PCR products from
micro-beads carrying 3' end of up-regulated cDNA. THP-1
cells induced with 100 nM PMA in DMSO."
BASE COUNT 27 a 19 c 14 g 27 t
ORIGIN
Query Match 3.9%; Score 87; DB 10; Length 87;

Fax: 216 361 9596
Email: scaine@athersys.com
High quality sequence stop: 441.
Location/Qualifiers
1..752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT 246 a 147 c 132 g 225 t
ORIGIN
Query Match 4.6%; Score 102; DB 12; Length 752;
Best Local Similarity 100.0%; Pred. No. 8.2e-21;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1983 CCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGGTCACTTCTACAACATAT 2042
Db 278 CCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGGTCACTTCTACAACATAT 219
QY 2043 GTTAGCACGGCAAGAACCTTCAATTTGAAGACTGAGATTTT 2084
Db 218 GTTAGCACGGCAAGAACCTTCAATTTGAAGACTGAGATTTT 177
RESULT 15
AW059922/c
LOCUS LF1b04.yg UPC15 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW059922
ACCESSION AW059922
VERSION AW059922.1 GI:6652244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 87)
AUTHORS Brenner,S., Williams,S.R., Vermass,E.H., Storck,T., Moon,K.,
McCollum,C., Mao,J.I., Kirchner,J.J., Eletr,S., DuBridge,R.B.,
Burcham,T. and Albrecht,G.
TITLE In vitro cloning of complex mixtures of DNA on microbeads: Physical
separation of differentially expressed cDNAs
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1665-1670 (2000)
MEDLINE 20144098
COMMENT Contact: Burcham TS
LYNX Therapeutics, Inc.
25861 Industrial Blvd., Hayward, CA 94545, USA
Tel: 510 670 9338
Fax: 510 670 9302
Email: timbelynxgen.com
Sequence obtained from LYNX Therapeutics Megasort technology.
Collected from the up-regulated gate.
High quality sequence stop: 87.
Location/Qualifiers
1..87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UPC15"
/cell_type="monocytic leukemia"
/note="Vector: pCR2.1; Cloning of PCR products from
micro-beads carrying 3' end of up-regulated cDNA. THP-1
cells induced with 100 nM PMA in DMSO."
BASE COUNT 27 a 19 c 14 g 27 t
ORIGIN
Query Match 3.9%; Score 87; DB 10; Length 87;

Best Local Similarity 100.0%; Pred. No. 1.4e-15;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1749 GATCTTGGCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTTCATTTCTGAA 1808
 |||||
Db 87 GATCTTGGCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTTCATTTCTGAA 28
 |||||
QY 1809 AGTCTAGAGAATTACAACTTGGTGAT 1835
 |||||
Db 27 AGTCTAGAGAATTACAACTTGGTGAT 1

Search completed: April 16, 2003, 14:10:03
Job time : 3558 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 06:13:54 ; Search time 93 Seconds
(without alignments)
7383.325 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	366.4	16.4	2072	2	US-09-073-362-2
2	366.4	16.4	2072	2	US-09-243-920-2
3	211.4	9.4	1011	2	US-08-825-781-2
4	67.4	3.0	7218	1	US-08-232-463-14
5	51.6	2.3	62804	4	US-09-800-960-3
6	50	2.2	152331	3	US-09-128-155-16
7	44.8	2.0	6124	4	US-08-213-419B-3
8	43.6	1.9	5375	3	US-08-757-223-7
9	43.4	1.9	893	4	US-09-370-838-166
10	43.4	1.9	40000	4	US-09-780-049-18
11	43	1.9	99500	4	US-09-798-096-10
12	43	1.9	162450	4	US-09-345-882-1
13	42.8	1.9	70000	4	US-09-851-896-3
14	42.8	1.9	72604	4	US-09-268-992-7
15	42.8	1.9	72604	4	US-09-657-474-7
16	42.6	1.9	3224	4	US-08-965-729A-2
17	42.6	1.9	12047	2	US-09-022-461-1
18	42.6	1.9	12047	4	US-09-033-556-3
19	42.4	1.9	72604	4	US-09-268-992-7
20	42.4	1.9	72604	4	US-09-657-474-7
21	42.2	1.9	1825	4	US-09-461-697-75
22	42.2	1.9	7218	1	US-08-232-463-14
23	41.8	1.9	301	2	US-08-332-766A-23
24	41.8	1.9	22481	4	US-08-367-841A-43
25	41.8	1.9	22481	5	PCT-US95-07201-43
26	41.8	1.9	22484	4	US-09-875-223-2
27	41.6	1.9	55827	4	US-09-813-133A-3

28	41.4	1.8	471	4	US-09-018-584A-6	Sequence 6, Appli
C 29	41.4	1.8	81001	4	US-09-750-580-1	Sequence 1, Appli
C 30	41.2	1.8	516	4	US-09-018-584A-24	Sequence 24, Appli
C 31	41	1.8	6405	4	US-09-281-481A-18	Sequence 18, Appli
32	41	1.8	17327	1	US-07-906-871-15	Sequence 15, Appli
33	41	1.8	38564	4	US-09-734-673-3	Sequence 3, Appli
C 34	41	1.8	111282	4	US-09-754-250-3	Sequence 3, Appli
35	40.8	1.8	87350	3	US-08-781-891-79	Sequence 79, Appli
36	40.8	1.8	87543	4	US-09-791-211-3	Sequence 3, Appli
37	40.6	1.8	45546	4	US-09-146-053-6	Sequence 6, Appli
38	40.4	1.8	2821	2	US-08-680-395-6	Sequence 6, Appli
39	40.2	1.8	4820	4	US-08-961-527-19	Sequence 19, Appli
C 40	40.2	1.8	12597	4	US-09-705-299-12	Sequence 12, Appli
41	40.2	1.8	14581	4	US-08-520-373D-4	Sequence 4, Appli
C 42	40.2	1.8	28720	4	US-09-341-587-7	Sequence 7, Appli
43	40	1.8	1001	4	US-09-641-638-401	Sequence 401, Appli
C 44	40	1.8	8453	4	US-09-167-681-45	Sequence 45, Appli
45	39.8	1.8	162450	4	US-09-345-882-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-073-362-2
; Sequence 2, Application US/09073362
; Patent No. 5942399
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2072 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: ESOGTUT02
; CLONE: 2667831
US-09-073-362-2
Query Match 16.4%; Score 366.4; DB 2; Length 2072;
Best Local Similarity 54.6%; Pred. No. 2.2e-92;

Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;

QY 300 CGCTTTTTCAGGAAGAGACGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAGAGGAAAGT 359
Db 192 CTCCCTTTGGGTGATGGGCGACCCAGGCGGAGGAGGTGAAGCTGAAGAGGAGAT 251

QY 360 CACTTTACTGAGGGAGTCTCCATATATCATTTGGACCATCATTTGGAGCAGGAATCTTCAT 419
Db 252 CTCACCTGCTTAACGGCGTGTGCTGATTTGGGGAACATGATCGGCTCAGGCACTTTTGT 311

QY 420 CTCTCTAAGGGCGTGTCCAGAACACGGGCGAGCGTGGGCATGTCTCTGACCATCTGGAC 479
Db 312 TTCCCCCAAGGGTGTGCTCATATACAGTGCCTCTCTTGGTCTCTCTCTGTCATCTGGGC 371

QY 480 GGTGTGTGGGCTCTGTCACTATTGAGCTTTGCTTATGCTGAATTTGGGAACAACATAT 539
Db 372 TGTCCGGGGCCTCTTCTCCGCTTTGGGGCCCTTTGTTATGCGGAACCTGGGCACCACTAT 431

QY 540 AAAGAAATCTGGAGGTCAATTACACATATATTTTGGGAAGTCTTTGGTCCATTACAGCTTT 599
Db 432 TAAGAAATCTGGGGCCAGCTATGCTATATCTTGGAGGCTTTGGAGGATTCCTTGTCTTT 491

QY 600 TGTACGAGTCTGGGTGGAACCTCTCATATAATACGCGCTGACGCTACTGCTGTGATATCCCT 659
Db 492 CATCAGACTCTGGACCTCCCTGCTCATCATTTGAGCCCAACAGCCAGGCACTCATTTGCCAT 551

QY 660 GGCATTTGGACGCTACATTTCTGGRAACCATTTTTTATTCATGTGAATCCCTGAACCTGC 719
Db 552 CACCTTTGCCAACTACATGGTACAGCCTCTCTTCCCGAGCTGCTTCGCCCTTTATGCTGC 611

QY 720 GATCAAGCTCATTAACAGCTGTGGGCATAACTGTAGTGTGCTTAAATAGCATGAGTGT 779
Db 612 CAGCCGCTGCTGGCTGCTGCTGCTGCTTAACTTCACTTAACTGAGCTGCTGCTGCTGCT 671

QY 780 CAGCTGGAGCGCCGATCCAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGAT 839
Db 672 CAAATGGGAACCCCTGGTACAAGATATTTTCACTTAACTTAACTGATTTGGCACTGATCGC 731

QY 840 AATTATAGTCCCTGGAGTTATGCAAGTAAATTAAGGTCAAACGCAAGCACTTTAAAGACGC 899
Db 732 GGTCACTGTTGAGGCACTTTGAGACTTGGCCAGGAGGCTTACTACTATTTTGAGAAATTC 791

QY 900 CTTTTCAGGAAGAGATTAAGTATTAAGCGGTTGCCACTGGCTTTTATTATTATGGAATGA 959
Db 792 CTTTGGAGGTTTCATCATTTTGCAGTGGGTGACATTTGCCCTGGCACTGTACTCAGCTCTGTT 851

QY 960 TGCATATGCTGGCTGTTTACCTCACTTTGTTACTGAAGAGTGAAGAAACCCCTGAAA 1019
Db 852 CTCCTACTCAGGCTGGGACACCCCTCACTATGTCACTGAAGAGATCAAGAAATCCTGAGAG 911

QY 1020 AACCATTTCCCTTGCATATGATATCCATGGCCATTTGTCAACCATTTGGCTATGCTGAC 1079
Db 912 GAACCTGCCCCCTCTCCATTTGGCATCTCCATGCCCCATTTGTCAACCATCATATATCTTGAC 971

QY 1080 AAATGTGGCCTACTTTACGACCACTTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGC 1139
Db 972 CAATGTGGCCTATATATGCTGCTAGACATGAGAGACATCTTGGCCAGTGTGCTGTTGC 1031

QY 1140 AGTGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTTAGCAGTTCGGATCTTTGTTGC 1199
Db 1032 TGTGACTTTTGCAGATCAGATATTTTGGAAATATTTAACTGGATATTTCCACTGTCAGTTGC 1091

QY 1200 CCTCTCTGCTTTGGCTCCATGAACGGTGTGTTGCTGTCTCCAGGTTATTTCTATGT 1259
Db 1092 ATTATCTGTTTGGTGGCCCTCAATGCCCTCCATTTGCTGCTGCTTCTAGGCTTTTCTTTGT 1151

QY 1260 TCGCTCTGAGAGGCTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACAC 1319
Db 1152 GGGCTCAAGAGAAGGCGCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGGTTTAC 1211

QY 1320 TCCTCTACAGCTGTATTTTGGACCCCTTTGACCAATGATAATGCTCTCTCTGGAGA 1379
Db 1212 ACCAGTGCCTTCTCTGCTCTTCAATGGTATCATGGCATTTGATCTACTTGTGCGTGAAGA 1271

QY 1380 CCTCGACAGTCTTTTGAATTTCTCAGTTTGGCCAGGTGGCTTTTATTGGGCTGGCAGT 1439
Db 1272 CATCTTCCAGCTCATTAACACTACTACAGCTTCAAGCTACTGGTCTTTTGTGGGCTTTCTAT 1331

QY 1440 TGCTGGGCTGATTTATCTTCGATACAAAATGCCCAGATATGCATCGTCTCTTCAAGGTGCC 1499
Db 1332 TGTGGTCAAGCTTTATCTGCGCTGGAAGAGCCTGATCGACCTCGTCCCTCAAGCTCAG 1391

QY 1500 ACTGTTCAFCACAGCTTTGTTTTCCTTTCACATGCCTCTTTCATGTTGCCCTTTCCCTCTA 1559
Db 1392 CGTTTCTTCCCGATTTGTTCTGCTGCTCTGCACCATCTTCTGTTGGTGTGTTCCACTTA 1451

QY 1560 TTCGGACCATTTAGTACAGGATTTGGCTTCGTCACTCACTGAGTGGAGTCCCTTGGTA 1619
Db 1452 CAGTGATACATCAACTCCCTCATCGGCATTTGCCATTTGCCCTCTCAGGCTGCCCCTTTTA 1511

QY 1620 TTATCTCTTTATTATA 1635
Db 1512 CTTCCTCATCATCAGA 1527

RESULT 2
US-09-243-920-2
; Sequence 2, Application US/09243920
; Patent No. 5981242
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,920
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/073,362
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2072 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: ESOGTUT02
; CLONE: 2667831
US-09-243-920-2

Query Match 16.4%; Score 366.4; DB 2; Length 2072;
Best Local Similarity 54.6%; Pred. No. 2.2e-92;
Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;

Db 21 CCCTGGCACTGTACTCAGCTCTGTTCTCCTACTCAGGCTGGGACACCTCAACTATGTCA 80
QY 995 CTGAAGAGTAGAAAACCCCTGAAAAAACCACTTCCCTTGCATATATATATCCATGGCCA 1054
Db 81 CTGAAGAGATCAAGAATCCTGAGAGAACCTGCCCTCTCCATGGCATCTCCATGGCCA 140
QY 1055 TTGTACCAATGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCAATTAATGCTGAGG 1114
Db 141 TTGTACCAATCATCTATATCTTGACCAATGTGGCCTATTATCTGTGCTAGACATGAGAG 200
QY 1115 AGCTGTGCTTTCAAATGCAGTGGGAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCT 1174
Db 201 ACATCTTGGCCAGTGATGCTGTTGCTGTGACTTTTGCAGATCAGATATTGGAATATTTA 260
QY 1175 CATTAGAGTTCCGATCTTTGTTGGCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGT 1234
Db 261 ACTGGATAATCCCACTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
QY 1235 TTGCTGTCTCCAGTTATTCTATGTGCGTCTCGAGAGGGTCACTTCCAGAAATCCCTCT 1294
Db 321 TGGCTGCTTCTAGGCTTTCTTTGCGGCTCAAGAGAGGCCCATCTCCCTGATGCCATCT 380
QY 1295 CCATGATTCATGTCCGCAAGCACACTCTCTACCAAGCTGTTATTGTTTGCACCCCTTTGA 1354
Db 381 GCATGATCCATGTTGAGCGGTTACACCAAGTGCTTCTCTGCTCTTCAATGGTATCATGG 440
QY 1355 CAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCTCAGTTTGGCCA 1414
Db 441 CATTGATCTACTGTGCGTGAAGACATCTTCCAGCTCATTAATACTACTACAGCTTCAGCT 500
QY 1415 GGTGGCTTTTATTGGGCTGGCAGTGTGCTGGGCTGATTTATCTCGATACAAATGCCAG 1474
Db 501 ACTGGTCTTGTGGGCTTTCTATGTGGGCTCAGCTTTATCTGCGCTGGAAGGAGCCTG 560
QY 1475 ATATGATTCGCTCTTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCTTCCATGCC 1534
Db 561 ATCGACCTCGTCCCTCAAGCTCAGCGTTTCTTCCGATTGCTCTGCTCTGCACTGCA 620
QY 1535 TCTTCATGTTGCCCTTTCCCTCTATTCGGACCCATTAGTACAGGATGGCTTCGTCA 1594
Db 621 TCTTCTGCTGGCTGTTCCACTTTACAGTGATATACTCAACTCCCTCATCGGCATGGCCA 680
QY 1595 TCACTCTGACTGGAGTCCCTGCGTATATCTCTTTATTATA 1635
Db 681 TTGCCCTCTCAGGCTGCCCTTTTACTTCTCTCATCATCAGA 721

RESULT 4

US-08-232-463-14

; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptZgpt-Fls
US-08-232-463-14

Query Match 3.0%; Score 67.4; DB 1; Length 7218;

Best Local Similarity 2.1%; Pred. No. 1.3e-08;

Matches 8; Conservative 235; Mismatches 136; Indels 0; Gaps 0;

QY 1184 TTCCGATCTTGTGCTCCTCTCCTGCTTGGCTCCATGAACGGTGGTGTGCTGTCT 1243

Db 1060 TTGCGATFYVV 1119

QY 1244 CCAGGTATTCTATGTTGCGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATC 1303

Db 1120 YVV 1179

QY 1304 ATGTCCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTGCACCCCTTGACAATGATA 1363

Db 1180 YVV 1239

QY 1364 TGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCAGGCTGCTTT 1423

Db 1240 YVV 1299

QY 1424 TTATTGGCTGGCAGTGTGCTGGGCTGATTTATCTCGATACAAATGCCAGATATCATC 1483

Db 1300 YVV 1359

QY 1484 GTCCTTTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCTTCCATGCTCTTTCATGG 1543

Db 1360 YVV 1419

QY 1544 TTGCCCTTCCCTCTATTTC 1562

Db 1420 YVV 1438

RESULT 5

US-09-800-960-3/c

; Sequence 3, Application US/09800960

; Patent No. 6387677

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001158

; CURRENT APPLICATION NUMBER: US/09/800,960

; CURRENT FILING DATE: 2001-03-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 62804

US-09-864-761-48172

Alignment Scores:

Pred. No.:	142	Length:	34
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	10	Gaps:	0

US-09-667-170A-440 (1-2239) x US-09-864-761-48172 (1-34)

QY 70 GAAAAAGAAAGAAAGAAAGAG 93

Db 24 GluLysArgLysLysGluLysGlu 31

Search completed: April 16, 2003, 16:57:06
Job time : 48 secs

Pred. No.: 1.19 Length: 523
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.37% Indels: 0
 DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-10-163-866-47 (1-523)

QY 409 GGAATCTTCATCTCTCCTAAGGCGTGCTC 438

Db 57 GlyilePheileserProLysGlyValLeu 66

RESULT 11

US-09-739-907-99
 ; Sequence 99, Application US/09739907
 ; Patent No. US20010012889A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 36 Human Secreted Proteins
 ; FILE REFERENCE: P2022P1
 ; CURRENT APPLICATION NUMBER: US/09/739,907
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 09/348,457
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: 60/070,567
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,692
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,704
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,658
 ; PRIOR FILING DATE: 1998-01-07
 ; NUMBER OF SEQ ID NOS: 196
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 99
 ; LENGTH: 132
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (132)
 ; OTHER INFORMATION: Xaa equals stop translation
 US-09-739-907-99

Alignment Scores:
 Pred. No.: 13 Length: 132
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.23% Indels: 0
 DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-739-907-99 (1-132)

QY 1993 TTAGGAGAAAGACTAGACAATTACTA 2019

Db 34 LeuGlyGluLysThrArgGlnLeu 42

RESULT 12

US-09-739-907-87
 ; Sequence 87, Application US/09739907
 ; Patent No. US20010012889A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 36 Human Secreted Proteins
 ; FILE REFERENCE: P2022P1
 ; CURRENT APPLICATION NUMBER: US/09/739,907
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 09/348,457
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: 60/070,567
 ; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,692
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,704
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,658
 ; PRIOR FILING DATE: 1998-01-07
 ; NUMBER OF SEQ ID NOS: 196
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 87
 ; LENGTH: 172
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (170)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-739-907-87

Alignment Scores:
 Pred. No.: 12.4 Length: 172
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.23% Indels: 0
 DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-739-907-87 (1-172)

QY 1993 TTAGGAGAAAGACTAGACAATTACTA 2019

Db 34 LeuGlyGluLysThrArgGlnLeu 42

RESULT 13

US-09-739-907-191
 ; Sequence 191, Application US/09739907
 ; Patent No. US20010012889A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 36 Human Secreted Proteins
 ; FILE REFERENCE: P2022P1
 ; CURRENT APPLICATION NUMBER: US/09/739,907
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 09/348,457
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: 60/070,567
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,692
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,704
 ; PRIOR FILING DATE: 1998-01-07
 ; PRIOR APPLICATION NUMBER: 60/070,658
 ; PRIOR FILING DATE: 1998-01-07
 ; NUMBER OF SEQ ID NOS: 196
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 191
 ; LENGTH: 182
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (180)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-739-907-191

Alignment Scores:
 Pred. No.: 12.3 Length: 182
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.23% Indels: 0
 DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-739-907-191 (1-182)

Thu Apr 17 09:48:31 2003

GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-41

Alignment Scores:
Pred. No.: 0.016 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-10-163-866-41 (1-511)

Qy 517 TATGCTGAATTGGGAACAACACTATAAAGAAATCTGGA 552
|||||
Db 90 TyrAlaGluLeuGlyThrThrIleLysSerGly 101

RESULT 8

US-10-163-866-42
; Sequence 42, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-42

Alignment Scores:
Pred. No.: 0.016 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-10-163-866-42 (1-511)
Qy 517 TATGCTGAATTGGGAACAACACTATAAAGAAATCTGGA 552
|||||
Db 90 TyrAlaGluLeuGlyThrThrIleLysSerGly 101

RESULT 9

US-09-815-923-16
; Sequence 16, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
; APPLICANT: Gill, Sarjeet S.
; APPLICANT: Ross, Linda S.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A
; FILE REFERENCE: 023070-093800US
; CURRENT APPLICATION NUMBER: US/09/815,923
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Aedes aegypti
; FEATURE:
; OTHER INFORMATION: LAT transporter
US-09-815-923-16

Alignment Scores:
Pred. No.: 0.138 Length: 517
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.50% Indels: 0
DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-815-923-16 (1-517)

Qy 409 GGAATCTTCATCTCTCTAAGGCGTGCTCCAG 441
|||||
Db 74 GlyllePheIleSerProLysGlyValLeuGln 84

RESULT 10

US-10-163-866-47
; Sequence 47, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-47

Alignment Scores:


```
RESULT 4
US-09-738-973-586
; Sequence 586, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 586
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-973-586
```

```
Alignment Scores:
Pred. No.: 3.17e-78 Length: 97
Score: 93.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.70% Indels: 0
DB: 10 Gaps: 0
```

US-09-667-170A-440 (1-2239) x US-09-738-973-586 (1-97)

```
QY 2 GAGTTGAAGTGAGCAGATCATGCCAGCCCTGGGTGACAGTGAGACTGTCTCAAAACA 61
Db 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20
QY 62 GAATTAAGGAAAAAGAAAAAGAAAAAGAGAGAGAGGAAATTCAGGCCAATTGTGGC 121
Db 21 GluLeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGly 40
QY 122 ATAGATTTTATCATATCTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTG 181
Db 41 IleAspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisHisTrpIleGln 60
QY 182 GAAAGCCTGTTGTGTCACCATCTCCAAAGAGAGGTACCTGCAGGAGAAATGTTAACGGGA 241
Db 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80
QY 242 GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCAGG 280
Db 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArg 93
```

```
RESULT 5
US-09-854-133-587
; Sequence 587, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 587
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-973-587
```

```
FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 587
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-133-587

Alignment Scores:
Pred. No.: 5.23e-06 Length: 16
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.19% Indels: 0
DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-854-133-587 (1-16)

QY 104 TTCCAGGCCCAATTGTGGCATAGATTTTATCATATTTCTGGATTTTGG 151
Db 1 PheGlnAlaAsnCysGlyIleAspPheIleIlePheTrpIlePheTrp 16
```

```
RESULT 6
US-09-738-973-587
; Sequence 587, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 587
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-973-587
```

```
Alignment Scores:
Pred. No.: 5.23e-06 Length: 16
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.19% Indels: 0
DB: 10 Gaps: 0
```

US-09-667-170A-440 (1-2239) x US-09-738-973-587 (1-16)

```
QY 104 TTCCAGGCCCAATTGTGGCATAGATTTTATCATATTTCTGGATTTTGG 151
Db 1 PheGlnAlaAsnCysGlyIleAspPheIleIlePheTrpIlePheTrp 16
```

```
RESULT 7
US-10-163-866-41
; Sequence 41, Application US/10163866
; Publication No. US20030027188A1
```

Db 36 GluLysValGlnLeuLysArgLysValThrLeuLeuArgGlyValSerIleIleGly 55
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCFAAGGCGTGTCCAGAACACGGGCAGC 453
Db 56 ThrIleIleGlyAlaGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer 75
QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCCCTGTCTACATATTGGAGCTTTG 513
Db 76 ValGlyMetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeu 95
QY 514 TCTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCAATACACATATATTTTG 573
Db 96 SerTyrAlaGluLeuGlyThrThrIleLysLysSerGlyGlyHisTyrThrTyrIleLeu 115
QY 574 GAAGTCTTTGGTCCATTACACGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATAACGC 633
Db 116 GluValPheGlyProLeuProAlaPheValArgValTrpValGluLeuLeuIleIleArg 135
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATTTCTGGAACCATTTT 693
Db 136 ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProphePhe 155
QY 694 ATTCAATGTGAAATCCCTGAACCTTGGCATCAAGCTCATACAGCTGTGGGCATACTGTA 753
Db 156 IleGlnCysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrVal 175
QY 754 GTGATGGTCTAAATAGCATGAGTGTACAGTGGAGCGCCGCGATCCAGATTTTCTTAACC 813
Db 176 ValMetValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThr 195
QY 814 TTTTGCAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAGCTAATTA 873
Db 196 PheCysLysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLys 215
QY 874 GGTCAAACGCGAAGCTTTAAAGACGCTTTTCAGGAAGAGATCAAGTATTACGCGGTG 933
Db 216 GlyGlnThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerIleThrArgLeu 235
QY 934 CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGTGGTGTACCTCAACTTTGTT 993
Db 236 ProLeuAlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheVal 255
QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAAACCATTTCCCTTGCATATGTATATCCATGGCC 1053
Db 256 ThrGluGluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAla 275
QY 1054 ATTGTCACCATTTGCTATGTCTGACAAATGTGGCCTACTTTACGACCATTAATGCTGAG 1113
Db 276 IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrIleAsnAlaGlu 295
QY 1114 GAGTGTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 296 GluLeuLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe 315
QY 1174 TCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCCATGAACGGTGTG 1233
Db 316 SerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGlyVal 335
QY 1234 TTTGCTGTCTCCAGGTATTCTATGTTGGCTGTCGAGAGGGTCACCTTCCAGAAATCCTC 1293
Db 336 PheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeu 355
QY 1294 TCCATGATTTCATGTCGCAAGCACACTCCTCTACAGCTGTTATTGTTTGCACCCCTTG 1353
Db 356 SerMetIleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeu 375
QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGC 1413
Db 376 ThrMetIleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAla 395
QY 1414 AGGTGGCTTTTATTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 396 ArgTrpLeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysPro 415

QY 1474 GATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATCCAGCTTTTGTCTTCCATCATGC 1533
Db 416 AspMetHisArgPropheLysValProLeuPheIleProAlaLeuPheSerPheThrCys 435
QY 1534 CTCTTCATGGTTGCCCTTCCCTCTATTCGGACCCCATTTAGTACAGGGATTGGCTTCGTC 1593
Db 436 LeuPheMetValAlaLeuSerLeuTyrSerAspPropheSerThrGlyIleGlyPheVal 455
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATATCTCTTTATATATGGGACAAAGAACCCAGG 1653
Db 456 IleThrLeuThrGlyValProAlaTyrTyrLeuPheIleIleTrpAspLysLysProArg 475
QY 1654 TGGTTTGAATAATGTGAGAGAAATAAACAGAACATTAACAATAATACTGGAAGTTGTA 1713
Db 476 TrpPheArgIleMetSerGluLysIleThrArgThrLeuGlnIleIleLeuGluValVal 495
QY 1714 CCAGAAGAAGATAAGTTA 1731
Db 496 ProGluGluAspLysLeu 501
RESULT 3
US-09-854-133-586
; Sequence 586, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 586
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-133-586
Alignment Scores:
Pred. No.: 3,17e-78 Length: 97
Score: 93.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.70% Indels: 0
DB: 9 Gaps: 0
US-09-667-170A-440 (1-2239) x US-09-854-133-586 (1-97)
QY 2 GAGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAACA 61
Db 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20
QY 62 GAATTAAGGAAAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCGCAATTTGTGGC 121
Db 21 GluLeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGly 40
QY 122 ATAGATTTTATCATATTCTGGATTTTGGATTTCTTTGTTTCTCATCATGATTGAG 181
Db 41 IleAspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisHisTrpIleGln 60
QY 182 GAAAGCCTGTTGTGTCCACCATCTCCAAAGGAGGTTTACCTGCAGGGAAATGTTAACGGGA 241
Db 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80
QY 242 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGCAGG 280
Db 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArg 93

Alignment Scores:

Pred. No.: 0 Length: 501
Score: 466.00 Matches: 466
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 63.66% Indels: 0
DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-10-163-866-48 (1-501)

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 36 GluLysValGlnLeuLysArgLysValThrLeuLeuArgGlyValSerIleIleGly 55
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGCGGTGCTCCAGAACACGGCAGC 453
Db 56 ThrIleIleGlyAlaGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer 75
QY 454 GTGGGCATGTCTCTGACCATCTGACGGGTGTGTGGGGTCCCTGTCACTATTGGAGCTTTG 513
Db 76 ValGlyMetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeu 95
QY 514 TCTTATGCTGAATTTGGAACTACTATAAGAAATCTGGAGGTCACTTACACATATATTTT 573
Db 96 SerTyrAlaGluLeuGlyThrIleLysLysSerGlyGlyHisTyrThrIleLeu 115
QY 574 GAAGTCTTTGGTCCATTACCACTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 633
Db 116 GluValPheGlyProLeuProAlaPheValArgValTrpValGluLeuIleIleArg 135
QY 634 CCTGCAGTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTGGAAACCATTTT 693
Db 136 ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProphePhe 155
QY 694 ATTCAATGTGAATCCCTGAATCTGCGATCGATCACTTACAGCTGTGGGCATAACTGTA 753
Db 156 IleGlnCysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrVal 175
QY 754 GTGATGGTCTCTAAATAGCATGAGTGTACGCTGGAGCGCCCGGATCCAGATTTTCTTAAC 813
Db 176 ValMetValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThr 195
QY 814 TTTTGAAGCTCACAGCAATCTGATAATATAGTCCCTGGAGTATGAGTATGAGTAAATAA 873
Db 196 PheCysLysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLys 215
QY 874 GGTCAACGCGAGACTTTAAAGACGCTTTTTCAGGAAGAGATCAAGTATTACGCGGTG 933
Db 216 GlyGlnThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerSerIleThrArgLeu 235
QY 934 CCACTGGCTTTTATGGAATGTATGATATGCTGGCTGCTTTTACCTCAACTTTGTT 993
Db 236 ProLeuAlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheVal 255
QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAAACCATTTCCCTTGCATATGTATATCCATGGCC 1053
Db 256 ThrGluGluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAla 275
QY 1054 ATTGTACCAATTTGGCTATGTGCTGACAAATGTGGCTTACTTTACGACCATTAATGCTGAG 1113
Db 276 IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrThrIleAsnAlaGlu 295
QY 1114 GAGCTGTGCTTTCAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 296 GluLeuLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe 315
QY 1174 TCATTAGCAGTCCGATCTTTTGTGCCCTCTCTCTGCTTTGGCTCCATGACGGTGGTGTG 1233
Db 316 SerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGlyVal 335
QY 1234 TTTGCTGTCTCCAGGTTATTTCTATGTTGCGTCTCGAGAGGGTCACTTCCAGAAATCCTC 1293
Db 336 PheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeu 355

QY 1294 TCCATGATTTCATGTCGCGAAGACACACTCCTCTACCAGCTGTTATTGTTTGCACCCCTTG 1353
Db 356 SerMetIleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeu 375
QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGC 1413
Db 376 ThrMetIleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAla 395
QY 1414 AGGTGGCTTTTATTGGCTGGCAGTTGCTGGGCTGATTATCTTCGATACAAATGCCCA 1473
Db 396 ArgTrpLeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysPro 415
QY 1474 GATATGCATCGTCTTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCTCCTTCACATGC 1533
Db 416 AspMetHisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCys 435
QY 1534 CTCTTCATGGTTGCCCTTCCCTCTATTTCGACCCCATTTAGTACAGGGATTGGCTTCGTC 1593
Db 436 LeuPheMetValAlaLeuSerLeuTyrSerAspPropheSerThrGlyIleGlyPheVal 455
QY 1594 ATCACTCTGACTGAGTCCCTGCGTATTATCTCTTTATTATATGGACAAAGAACCCAGG 1653
Db 456 IleThrLeuThrGlyValProAlaTyrTyrLeuPheIleIleTrpAspLysLysProArg 475
QY 1654 TGGTTTAGAATAATGTACAGAGAAAATAACCAAGAACATTACAAATAATACTGGAGTTGTA 1713
Db 476 TrpPheArgIleMetSerGluLysIleThrArgThrLeuGlnIleIleLeuGluValVal 495
QY 1714 CCAGAAGAAGATAAGTTA 1731
Db 496 ProGluGluAspLysLeu 501

RESULT 2

US-10-163-866-49
; Sequence 49, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-49

Alignment Scores:

Pred. No.: 0 Length: 501
Score: 466.00 Matches: 466
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 63.66% Indels: 0
DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-10-163-866-49 (1-501)

QY 334 .GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
|||||

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:40:19 ; Search time 42 Seconds
(without alignments)
8061.880 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 732
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 288829 seqs, 75613885 residues

Word size: 1
Total number of hits satisfying chosen parameters: 541640

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09667170/runat_04042003_091003_21265/app_query.fasta_1.2375
-DB=Published_Applications_AA -Qfmt=fastan -SUFFIX=olin2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09667170@cgn_1_1_15_runat_04042003_091003_21265
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	466	63.7	501	9	US-10-163-866-48
2	466	63.7	501	9	US-10-163-866-49
3	93	12.7	97	9	US-09-854-133-586
4	93	12.7	97	10	US-09-738-973-586

5	16	2.2	16	9	US-09-854-133-587	Sequence 587, App
6	16	2.2	16	10	US-09-738-973-587	Sequence 587, App
7	12	1.6	511	9	US-10-163-866-41	Sequence 41, Appl
8	12	1.6	511	9	US-10-163-866-42	Sequence 42, Appl
9	11	1.5	517	9	US-09-815-923-16	Sequence 16, Appl
10	10	1.4	523	9	US-10-163-866-47	Sequence 47, Appl
11	9	1.2	132	10	US-09-739-907-99	Sequence 99, Appl
12	9	1.2	172	10	US-09-739-907-87	Sequence 87, Appl
13	9	1.2	182	10	US-09-739-907-191	Sequence 191, Appl
14	8	1.1	34	10	US-09-864-761-41510	Sequence 41510, A
15	8	1.1	34	10	US-09-864-761-48172	Sequence 48172, A
16	8	1.1	49	10	US-09-864-761-42434	Sequence 42434, A
17	8	1.1	107	9	US-10-083-357-700	Sequence 700, App
18	8	1.1	107	9	US-10-083-357-1292	Sequence 1292, App
19	8	1.1	152	10	US-09-925-301-900	Sequence 900, App
20	8	1.1	179	10	US-09-864-761-43216	Sequence 43216, A
21	8	1.1	180	9	US-10-163-866-50	Sequence 50, Appl
22	8	1.1	414	10	US-09-925-297-747	Sequence 747, App
23	8	1.1	507	9	US-10-163-866-38	Sequence 38, Appl
24	8	1.1	507	9	US-10-163-866-39	Sequence 39, Appl
25	8	1.1	507	9	US-10-163-866-54	Sequence 54, Appl
26	8	1.1	515	9	US-10-163-866-40	Sequence 40, Appl
27	8	1.1	524	9	US-10-163-866-37	Sequence 37, Appl
28	8	1.1	535	9	US-10-163-866-43	Sequence 43, Appl
29	8	1.1	535	9	US-10-163-866-44	Sequence 44, Appl
30	8	1.1	594	9	US-10-028-072-10	Sequence 10, Appl
31	8	1.1	594	9	US-10-121-049-10	Sequence 10, Appl
32	8	1.1	594	9	US-10-123-904-10	Sequence 10, Appl
33	8	1.1	594	9	US-10-140-470-10	Sequence 10, Appl
34	8	1.1	594	9	US-10-175-746-10	Sequence 10, Appl
35	8	1.1	594	9	US-10-176-918-10	Sequence 10, Appl
36	8	1.1	594	9	US-10-176-921-10	Sequence 10, Appl
37	8	1.1	594	9	US-10-137-865-10	Sequence 10, Appl
38	8	1.1	594	9	US-10-140-474-10	Sequence 10, Appl
39	8	1.1	594	9	US-10-142-431-10	Sequence 10, Appl
40	8	1.1	594	9	US-10-143-114-10	Sequence 10, Appl
41	8	1.1	594	9	US-10-140-002-10	Sequence 10, Appl
42	8	1.1	594	9	US-10-142-419-10	Sequence 10, Appl
43	8	1.1	594	9	US-10-123-262-10	Sequence 10, Appl
44	8	1.1	594	9	US-10-142-423-10	Sequence 10, Appl
45	8	1.1	594	9	US-10-121-050-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-163-866-48
; Sequence 48, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-48

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Alignment Scores:
Pred. No.: 16.6 Length: 2227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 3 Gaps: 0

US-09-667-170A-440 (1-2239) x US-08-475-886-6 (1-2227)

QY 990 AAAGTTGAGGTAAACCCAGCCAGC 967
|||||
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 15
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Alignment Scores:
Pred. No.: 16.6 Length: 2227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x US-08-397-232-2 (1-2227)

QY 990 AAAGTTGAGGTAAACCCAGCCAGC 967
|||||
Db 1350 LysValGluValLysProAlaSer 1357

Search completed: April 16, 2003, 16:42:51
Job time : 33.5 secs

```
; TITLE OF INVENTION: AND VACCINES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04910
; FILING DATE: 21-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/231,526
; FILING DATE: 22-APR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4125PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 995 amino acids
; TYPE: amino acids
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; PCT-US95-04910-14

Alignment Scores:
Pred. No.: 18.3 Length: 995
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 5 Gaps: 0

US-09-667-170A-440 (1-2239) x PCT-US95-04910-14 (1-995)
QY 1727 TTATCTCTCTGGTACAACCTTC 1704
|||||
Db 434 LeuSerSerGlyThrThrSer 441

RESULT 12
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT

US-09-667-170A-440 (1-2239) x US-08-475-886-2 (1-2227)
QY 990 AAAGTTGAGGTAAACCCAGCCAGC 967
|||||
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 13
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT

US-09-667-170A-440 (1-2239) x US-08-475-886-4 (1-2227)
QY 990 AAAGTTGAGGTAAACCCAGCCAGC 967
|||||
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 14
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT

; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Alignment Scores:
Pred. No.: 16.6 Length: 2227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 3 Gaps: 0

US-09-667-170A-440 (1-2239) x US-08-475-886-2 (1-2227)
QY 990 AAAGTTGAGGTAAACCCAGCCAGC 967
|||||
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 13
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT

US-09-667-170A-440 (1-2239) x US-08-475-886-4 (1-2227)
QY 990 AAAGTTGAGGTAAACCCAGCCAGC 967
|||||
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 14
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
```

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%                Indels: 0
DB: 4                             Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-556-273-11 (1-786)

QY 766 AATAGCATGAGTGTCTGAGGAGC 789
Db 549 AsnSerMetSerValSerTrpSer 556

RESULT 9
US-09-012-710-10
; Sequence 10, Application US/09012710
; Patent No. 6087478
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; APPLICANT: Moarefi, Ismail
; APPLICANT: Darnell, Jr., James E.
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,710
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/012,710
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-194
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-012-710-10

Alignment Scores:
Pred. No.: 18.8      Length: 793
Score: 8.00          Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%                Indels: 0
DB: 3                 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-012-710-10 (1-793)

QY 766 AATAGCATGAGTGTCTGAGGAGC 789
Db 549 AsnSerMetSerValSerTrpSer 556

RESULT 10
US-09-556-273-10

```

```

; Sequence 10, Application US/09556273
; Patent No. 6312887
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; APPLICANT: Moarefi, Ismail
; APPLICANT: Darnell, Jr., James E.
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/556,273
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/012,710
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-194
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-556-273-10

Alignment Scores:
Pred. No.: 18.8      Length: 793
Score: 8.00          Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%                Indels: 0
DB: 4                 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-556-273-10 (1-793)

QY 766 AATAGCATGAGTGTCTGAGGAGC 789
Db 549 AsnSerMetSerValSerTrpSer 556

RESULT 11
PCT-US95-04910-14
; Sequence 14, Application PC/TUS9504910
; GENERAL INFORMATION:
; APPLICANT: The Government of the United
; APPLICANT: States of America as represented
; APPLICANT: by the Secretary, Department of
; APPLICANT: Health and Human Services
; TITLE OF INVENTION: ISOLATION AND
; TITLE OF INVENTION: CHARACTERIZATION OF A NOVEL PRIMATE T-CELL
; TITLE OF INVENTION: LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS
; TITLE OF INVENTION: OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS

```


US-09-071-035-306
; Sequence 306, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 306:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-306

Alignment Scores:
Pred. No.: 21.3 Length: 275
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-071-035-306 (1-275)

QY 437 AGCACGCCCTTAGGAGAGATGAAG 414
|||||
Db 125 SerThrProLeuGlyGluMetLys 132

RESULT 5
US-09-073-362-3
; Sequence 3, Application US/09073362
; Patent No. 5942399
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1665759
US-09-073-362-3

Alignment Scores:
Pred. No.: 19.8 Length: 515
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-073-362-3 (1-515)

QY 517 TATGCTGAATGGGAACAATA 540
|||||
Db 98 TyrAlaGluLeuGlyThrThrile 105

RESULT 6
US-09-243-920-3
; Sequence 3, Application US/09243920
; Patent No. 5981242
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,920
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/073,362

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run On: April 16, 2003, 16:33:19 ; Search time 18.5 Seconds
(without alignments)
7121.938 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 732
Sequence: 1 ggaggttgagtgagcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues
Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_091002_21223/app_query.fasta_1.2375
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=olin2p.ra -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_17_@runat_04042003_091002_21223 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	1.6	511	2	US-09-073-362-1
2	12	1.6	511	2	US-09-243-920-1
C 3	8	1.1	241	4	US-09-071-035-308
C 4	8	1.1	275	4	US-09-071-035-306
5	8	1.1	515	2	US-09-073-362-3
6	8	1.1	515	2	US-09-243-920-3
7	8	1.1	786	3	US-09-012-710-11
8	8	1.1	786	4	US-09-556-273-11
9	8	1.1	793	3	US-09-012-710-10
10	8	1.1	793	4	US-09-556-273-10
C 11	8	1.1	995	5	PCT-US95-04910-14
C 12	8	1.1	2227	3	US-08-475-886-2

C 13	8	1.1	2227	3	US-08-475-886-4	Sequence 4, Appli
C 14	8	1.1	2227	3	US-08-475-886-6	Sequence 6, Appli
C 15	8	1.1	2227	4	US-08-397-232-2	Sequence 2, Appli
C 16	8	1.1	2227	4	US-08-397-232-4	Sequence 4, Appli
C 17	8	1.1	2227	4	US-09-171-387-2	Sequence 2, Appli
C 18	8	1.1	2227	4	US-09-653-499-2	Sequence 2, Appli
C 19	8	1.1	2227	4	US-09-653-499-4	Sequence 4, Appli
C 20	8	1.1	2227	4	US-09-653-499-6	Sequence 6, Appli
C 21	7	1.0	12	2	US-08-406-330-44	Sequence 44, Appl
C 22	7	1.0	12	2	US-08-556-597-44	Sequence 44, Appl
C 23	7	1.0	18	1	US-07-972-032-8	Sequence 8, Appli
C 24	7	1.0	18	1	US-08-642-255-8	Sequence 8, Appli
C 25	7	1.0	21	4	US-08-817-548A-3	Sequence 3, Appli
C 26	7	1.0	22	3	US-08-940-095-156	Sequence 156, App
C 27	7	1.0	22	3	US-08-940-095-158	Sequence 158, App
C 28	7	1.0	22	3	US-08-940-095-166	Sequence 166, App
C 29	7	1.0	22	3	US-08-940-095-172	Sequence 172, App
C 30	7	1.0	22	3	US-08-940-095-177	Sequence 177, App
C 31	7	1.0	22	3	US-08-940-095-184	Sequence 184, App
C 32	7	1.0	22	3	US-08-940-093-156	Sequence 156, App
C 33	7	1.0	22	3	US-08-940-093-158	Sequence 158, App
C 34	7	1.0	22	3	US-08-940-093-166	Sequence 166, App
C 35	7	1.0	22	3	US-08-940-093-172	Sequence 172, App
C 36	7	1.0	22	3	US-08-940-093-177	Sequence 177, App
C 37	7	1.0	22	3	US-08-940-093-184	Sequence 184, App
C 38	7	1.0	22	3	US-08-940-096-156	Sequence 156, App
C 39	7	1.0	22	3	US-08-940-096-158	Sequence 158, App
C 40	7	1.0	22	3	US-08-940-096-166	Sequence 166, App
C 41	7	1.0	22	3	US-08-940-096-172	Sequence 172, App
C 42	7	1.0	22	3	US-08-940-096-177	Sequence 177, App
C 43	7	1.0	22	3	US-08-940-096-184	Sequence 184, App
C 44	7	1.0	22	4	US-09-465-719-156	Sequence 156, App
C 45	7	1.0	22	4	US-09-465-719-158	Sequence 158, App

ALIGNMENTS

RESULT 1
US-09-073-362-1
; Sequence 1, Application US/09073362
; Patent No. 5942399
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166


```

Db      1726 TGGAT 1730
|||||
RESULT 14
AB011263
LOCUS      1536 bp mRNA linear PRI 15-AUG-2000
DEFINITION Homo sapiens mRNA for MOP-2, complete cds.
ACCESSION AB011263
VERSION    AB011263.1 GI:9836571
KEYWORDS   MOP-2.
SOURCE     Homo sapiens monocyte mRNA.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1536)
AUTHORS    Takayama,K. and Yoshimoto,M.
TITLE      Molecular and Biological Characterization of a Novel Monocyte Amino
            Acid Permease, MOP-2
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1536)
AUTHORS    Takayama,K. and Yoshimoto,M.
TITLE      Direct Submission
JOURNAL    Submitted (23-FEB-1998) Kiyoshi Takayama, Taisho Pharmaceutical Co,
            Ltd., Molecular Biology Laboratory; Yoshino-cho 1-403, Ohmiya,
            Saitama 330, Japan (E-mail:s1677@cmm.taisho.co.jp,
            Tel:+81-48-663-1111, Fax:+81-48-652-7254)
FEATURES   Location/Qualifiers
            1..1536
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /cell_type="monocyte"
             1..1536
              /note="similar to a amino acid permease"
              /codon_start=1
              /product="MOP-2"
              /protein_id="BAB11849.1"
              /db_xref="GI:9836572"
              /translation="MVDSTYEYVASQPEVETSPLDGASPGPEQVKLKKEISLNGVC
                LIVGNMGISGIFVPKGVLIYSASFGLSLVIWAVGGFLSVFGALCYAELGTTIKSGA
                SYAILAEFGFLAFIRLWTSLLIIPTSQAIITFANYMVQPLFPSCFAPYAASRL
                LAAACICLLTFINCAYVKWCTLVQDIFTYAKVLALIAIVAGIVRLQGQASTHFENSF
                EGSSFAVGDIADIALYSALFSYSGWDTLNVTTEEIKPERNLPLSIGISMPVITIYL
                TNVAYTVLDMRDILASDAVAVTFADQIFGFNWIIPLSVALSFCGGNLNASIVAASRL
                FVGSREGHPDAICMIHVREFTPVPSLLFNIGIMALIYLCVEDIFQLINYVSFSYWF
                VGLSIVQLYLRWKEPDRPLKLSVFFPIVFCLCTFLVAVPLYSDTINSLGIAIA
                LSGLPFYFLIIRVPEHKRPLYLRRIVGSATRYLVLCMSVAEMDLEDGGEPMKORDP
                KSN"
BASE COUNT      303 a      420 c      364 g      449 t
ORIGIN
Query Match          16.4%; Score 366.4; DB 9; Length 1536;
Best Local Similarity 54.6%; Pred. No. 9.5e-76;
Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;
QY 300 CGCCTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGAGAAAAGTGCAAGTGAAGAGGAAAGT 359
      |||||
Db  51 CTCGCCCTTTGGGTGATGGGGCCAGCCCGAGCGGCGGAGCAGGTGAAGCTGAAGAAGGAGAT 110
      |||||
QY 360 CACTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCAT 419
      |||||
Db  111 CTCAC TGCTTAACGGCGTGTGCCTGATGTGGGGAACATGATCGGCTCAGGCATCTTTGT 170
      |||||
QY 420 CTCTCCTAAGGCGTGCTCCAGAACACCGGCGACGGTGGGCATGCTCTGTGACCATCTGGAC 479
      |||||
Db  171 TTCCCCCAAGGGTGTGCTCATATACAGTGCCTCTTTGGTCTCTCTCTGGTTCATCTGGGC 230
      |||||
QY 480 GG GTGTGGGGTCTGTGCTACTATTTGGAGCTTTGTCTTATGCTGAATGGGAACAACATAT 539
      |||||
Db  231 TGTCCGGGGCCCTCTTCTCCGTCTTTGGGGCCCTTTGTGTTATCGGAACTGGGCACCAACCAT 290
      |||||
QY 540 AAAGAAATCTGGAGGTCATTACACATATAATTTTGGGAAGTCTTTGGTCCATTACCAAGCTTT 599
      |||||

```


misc_feature 4137. .4527 /note="match to EST AI267688 (NID:g3886855) aq93cl0.xl"
misc_feature 4153. .4433 /note="match to EST C00538 (NID:g1432768)"
misc_feature 4233 /note="match to EST BI497113 (NID:g15336457)"
misc_feature 4359. .4746 /note="match to EST AA227760 (NID:g1849333) zr56a07.r1"
misc_feature 4404. .5178 /note="match to EST BG291471 (NID:g13049386)"
misc_feature 4412. .4619 /note="similar to Homo sapiens EST BF154977 (NID:g11050160)"
misc_feature 4657. .5028 /note="match to EST R20307 (NID:g774941) yg35a09.r1"
misc_feature 4727. .5195 /note="similar to Bos taurus EST AW669076 (NID:g7525590)"
misc_feature 4740. .5803 /note="similar to EST BM547723 (NID:g18781736)"
misc_feature 4859. .5265 /note="match to EST AW901105 (NID:g8065310)"
misc_feature 5002. .5687 /note="match to EST BG254117 (NID:g12763933)"
misc_feature 5004. .5213 /note="match to EST BF084219 (NID:g10878049)"
misc_feature 5074. .5447 /note="match to EST N80420 (NID:g1243121) yz92e06.r1"
misc_feature 5255. .6002 /note="match to EST BI869486 (NID:g16043159)"
misc_feature 5442. .6151 /note="match to EST AW961203 (NID:g8150887)"
misc_feature 5497. .5830 /note="match to EST AI540897 (NID:g4458270)"
misc_feature 5782. .6124 /note="match to EST AA309186 (NID:g1961511)"
misc_feature 5875. .6701 /note="match to EST BI334972 (NID:g15019629)"
misc_feature 5995. .6331 /note="match to EST AA471165 (NID:g2199022)"
misc_feature 6000. .6124 /note="match to EST AW966896 (NID:g8257130)"
misc_feature 6032. .6836 /note="similar to EST BM474943 (NID:g18523985)"
misc_feature 6140. .6518 /note="match to EST AW996897 (NID:g8257131)"
misc_feature 6444. .6885 /note="match to EST AA375957 (NID:g2028275)"
repeat_region 6878. .7187 /rpt_family="Alu"
misc_feature 6899. .7362 /note="match to EST AA488687 (NID:g2216118) ab38f03.s1"
misc_feature 6970. .7746 /note="match to EST BE739933 (NID:g10153837)"
misc_feature 7099. .7815 /note="match to EST BF381953 (NID:g11363256)"
misc_feature 7124. .7125 /note="match to EST AW996897 (NID:g8257131)"
misc_feature 7174. .7495 /note="match to EST BF927528 (NID:g12325656)"
misc_feature 7253. .7884 /note="similar to Homo sapiens EST BG499179 (NID:g13460696)"
misc_feature 7568. .7949 /note="match to EST AA307001 (NID:g1959330)"
misc_feature 7613. .8332 /note="similar to EST BM706727 (NID:g19019985)"
repeat_region 7836. .8011 /rpt_family="AcHobo"
misc_feature 8035. .8572

Query Match 23.6%; Score 529.2; DB 9; Length 64196;
Best Local Similarity 99.4%; Pred. No. 2.9e-114;
Matches 531; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1671 AGAGAAAATAACACAGAACATTACAAAATAATACTGGAAGTTGTACCAGAGAAGATAAGTT 1730
|||||
Db 11755 AGAGAAAATAACACAGAACATTACAAAATAATACTGGAAGTTGTACCAGAGAAGATAAGTT 11696
|||||
QY 1731 ATGAACATAATGGACTTGGAGATCTTGGCAATCTGCCAAAGGGGAGACACACAAATAGGGATT 1790
|||||
Db 11695 ATGAACATAATGGACTTGGAGATCTTGGCAATCTGCCAAAGGGGAGACACACAAATAGGGATT 11636
|||||
QY 1791 TTTACTTTCATTTTCTGAAAGTCTAGAGAAATTACAACTTTGGTGATAAAACAAAAGGAGTCA 1850
|||||
Db 11635 TTTACTTTCATTTTCTGAAAGTCTAGAGAAATTACAACTTTGGTGATAAAACAAAAGGAGTCA 11576
|||||
QY 1851 GTTATTTTATTTCATATATTTTAGCATATATTCGAACATAATTTCTAAGAAATTTAGTTATAA 1910
|||||
Db 11575 GTTATTTTATTTCATATATTTTAGCATATATTCGAACATAATTTCTAAGAAATTTAGTTATAA 11516
|||||
QY 1911 CTCTATGTAGTTATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTTGTAGT 1970
|||||
Db 11515 CTCTATGTAGTTATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTTGTAGT 11456
|||||
QY 1971 CTCTGATACCTACCTATTGGGTTAGGAGAAAAGACTAGACAAATTTACTATGTGTCATTC 2030
|||||
Db 11455 CTCTGATACCTACCTATTGGGTTAGGAGAAAAGACTAGACAAATTTACTATGTGTCATTC 11396
|||||
QY 2031 TCTACAACATATGTTAGCACGCAAGAAACCTTCAAAATTTGAAGACTGAGATTTTCTGTGA 2090
|||||
Db 11395 TCTACAACATATGTTAGCACGCAAGAAACCTTCAAAATTTGAAGACTGAGATTTTCTGTGA 11336
|||||
QY 2091 TATATGGGTTTTGTAAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAAGTGT 2150
|||||
Db 11335 TATATGGGTTTTGTAAAGATGGTTTACACACTATAGATGTCTATACTGTGAAAAGTGT 11276
|||||
QY 2151 TTCAATTCTGAAAAAAGCAATACATCATGATTATGGCAAGAGGAGAGAGGTA 2204
|||||
Db 11275 TTCAATTCTGAAAAAAGCAATACATCATGATTATGGCAAGAGGAGAGAGAGGTA 11222
|||||

RESULT 11
BC028216 3448 bp mRNA linear PRI 01-MAY-2002
LOCUS Homo sapiens, solute carrier family 7 (cationic amino acid transporter, y+ system), member 6, clone MGC:40002 IMAGE:5214340, mRNA, complete cds.
ACCESSION BC028216 GI:20381371
VERSION BC028216.1
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3448)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Laric, P., Legaspi, R., Maduro, Q.L.,


```

/gene="xCT"
/function="anionic amino acid transporter"
/codon_start=1
/evidence=experimental
/product="cystine/glutamate transporter"
/protein_id="BAA77220.1"
/db_xref="GI:4689081"
/translation="MVRKPVVATISKGGYLGQNMGRSLPMSGDOEPPEQKVVLLKKKI
TLRGVSIIGTVIGSGIFISPKGILONTGVSMSLVFWSACGVLSLFGALSVAELGT
SIKSSGSHYTYILEVEGPLLAFVRVWVELLVIRPGATAVISLAFGRYILEPFIQCEI
PELAIKLVAVGITVVMVNSTSVSWSARIQIFLFFCKLTAILIIVPGVIQLIKQGT
HHFKDAFSGRDTSLMGLPLAFYGYMYAYAGWFYLNFTIEVDNPEKTIPLAICISMAI
ITGVYVLTNVAYFTTISAEELIQSSAVAVTFSERLIGKFSVAVPIFVALSCFGSMNGG
VFAVSRLFYVASREGLHPEILSMIHVHKHTPLPAVIVLHPLTMVLMFSGDLYSLNPL
SFARWLEMLAVAGLIYLRYPKPDHMRPKVPLFIPALFSFTCLFMVLSLSDPFSF
GVGFLITLTGVPAYLYLFIVWDRKPKWFRRLSDRITRLQIILEVVPEDSKEL"
2216
polyA_site
/gene="xCT"
/note="34 a nucleotides"
BASE COUNT 542 a 526 c 509 g 639 t
ORIGIN
Query Match 54.0%; Score 1208.6; DB 10; Length 2216;
Best Local Similarity 79.1%; Pred. No. 1.3e-274;
Matches 1548; Conservative 0; Mismatches 324; Indels 85; Gaps 6;
QY 182 GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTAAACGGGA 241
D 182 GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTAAACGGGA 241
D 182 GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTAAACGGGA 241
QY 242 GGCTGCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGACGCCTTTTTCAGGAAGAGACG 301
D 242 GGCTGCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGACGCCTTTTTCAGGAAGAGACG 301
D 242 GGCTGCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGACGCCTTTTTCAGGAAGAGACG 301
QY 302 CCTTTTCAGGAAGAGACGCCTTTTTCAGGAAGAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
D 302 CCTTTTCAGGAAGAGACGCCTTTTTCAGGAAGAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
D 302 CCTTTTCAGGAAGAGACGCCTTTTTCAGGAAGAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
QY 425 -----GAGRAGGTAGTTCTGAAAAGAGAGATCA 452
D 425 -----GAGRAGGTAGTTCTGAAAAGAGAGATCA 452
D 425 -----GAGRAGGTAGTTCTGAAAAGAGAGATCA 452
QY 362 CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCATCT 421
D 362 CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCATCT 421
D 362 CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCATCT 421
QY 453 CTTTGTGAGGGGGGTCTCCATCATCATCGGCACCGCTCATCGGATCAGGCATCTTCATCT 512
D 453 CTTTGTGAGGGGGGTCTCCATCATCATCGGCACCGCTCATCGGATCAGGCATCTTCATCT 512
D 453 CTTTGTGAGGGGGGTCTCCATCATCATCGGCACCGCTCATCGGATCAGGCATCTTCATCT 512
QY 422 CTCCTAAGGGCGTGTCCAGAACACGGGCGAGCGTGGGATGCTCTGTACCATCTGGACGG 481
D 422 CTCCTAAGGGCGTGTCCAGAACACGGGCGAGCGTGGGATGCTCTGTACCATCTGGACGG 481
D 422 CTCCTAAGGGCGTGTCCAGAACACGGGCGAGCGTGGGATGCTCTGTACCATCTGGACGG 481
QY 513 CCCCCAAGGGCATACTCCAGAACACGGGCGAGCGTGGGATGCTCCCTGGTCTTCTGGTCTG 572
D 513 CCCCCAAGGGCATACTCCAGAACACGGGCGAGCGTGGGATGCTCCCTGGTCTTCTGGTCTG 572
D 513 CCCCCAAGGGCATACTCCAGAACACGGGCGAGCGTGGGATGCTCCCTGGTCTTCTGGTCTG 572
QY 482 TGTGTGGGGTCTGTGCTACTATTGGAGCTTTTGTCTTATGCTGAATTTGGGAACACTATAA 541
D 482 TGTGTGGGGTCTGTGCTACTATTGGAGCTTTTGTCTTATGCTGAATTTGGGAACACTATAA 541
D 482 TGTGTGGGGTCTGTGCTACTATTGGAGCTTTTGTCTTATGCTGAATTTGGGAACACTATAA 541
QY 573 CCGTGGAGTACTGTACATTTTGGAGCCCTGTCTCATATGAGAAATAGGTACAGCATAA 632
D 573 CCGTGGAGTACTGTACATTTTGGAGCCCTGTCTCATATGAGAAATAGGTACAGCATAA 632
D 573 CCGTGGAGTACTGTACATTTTGGAGCCCTGTCTCATATGAGAAATAGGTACAGCATAA 632
QY 542 AGAAATCTGGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAAGCTTTG 601
D 542 AGAAATCTGGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAAGCTTTG 601
D 542 AGAAATCTGGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAAGCTTTG 601
QY 633 AGAAATCTGGTGGTCAATACACATATATTTTGGAGGTCTTTGGTCTTTTGTGGCTTTG 692
D 633 AGAAATCTGGTGGTCAATACACATATATTTTGGAGGTCTTTGGTCTTTTGTGGCTTTG 692
D 633 AGAAATCTGGTGGTCAATACACATATATTTTGGAGGTCTTTGGTCTTTTGTGGCTTTG 692
QY 602 TACGAGTCTGGGTGGAACCTCCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661
D 602 TACGAGTCTGGGTGGAACCTCCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661
D 602 TACGAGTCTGGGTGGAACCTCCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661
QY 693 TTCGAGTCTGGGTGGAACCTGCTCGTAATACGCCCTGGAGCTACTGCTGTGATATCCCTGG 752
D 693 TTCGAGTCTGGGTGGAACCTGCTCGTAATACGCCCTGGAGCTACTGCTGTGATATCCCTGG 752
D 693 TTCGAGTCTGGGTGGAACCTGCTCGTAATACGCCCTGGAGCTACTGCTGTGATATCCCTGG 752
QY 662 CATTTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCGA 721
D 662 CATTTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCGA 721
D 662 CATTTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCGA 721
QY 753 CATTTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCGA 812
D 753 CATTTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCGA 812
D 753 CATTTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCGA 812
QY 722 TCAAGCTCAATTACAGCTGTGGGCAATACTGTAGTGTGCTCCTAAATAGCATGAGTGCA 781
D 722 TCAAGCTCAATTACAGCTGTGGGCAATACTGTAGTGTGCTCCTAAATAGCATGAGTGCA 781
D 722 TCAAGCTCAATTACAGCTGTGGGCAATACTGTAGTGTGCTCCTAAATAGCATGAGTGCA 781
QY 813 TCAAGCTCTGACAGCTGTGGGCAATCACTGTGTGTGATGCTGCTAAATAGCAGTGCTCA 872
D 813 TCAAGCTCTGACAGCTGTGGGCAATCACTGTGTGTGATGCTGCTAAATAGCAGTGCTCA 872
D 813 TCAAGCTCTGACAGCTGTGGGCAATCACTGTGTGTGATGCTGCTAAATAGCAGTGCTCA 872
QY 782 GCTGGAGCGCGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 841
D 782 GCTGGAGCGCGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 841
D 782 GCTGGAGCGCGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 841
QY 873 GCTGGAGTGGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 932
D 873 GCTGGAGTGGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 932
D 873 GCTGGAGTGGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 932
QY 842 TTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGCAACTTTAAAGACGCT 901
D 842 TTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGCAACTTTAAAGACGCT 901
D 842 TTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGCAACTTTAAAGACGCT 901
QY 933 TTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGCAACTTTAAAGATGCAT 992
D 933 TTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGCAACTTTAAAGATGCAT 992
D 933 TTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGCAACTTTAAAGATGCAT 992
```

```

QY 902 TTTCAGGAAGAGAGATTCAAGTATTACGGGTTGCCAGTGGCTTTTATTATGAATGTATG 961
D 902 TTTCAGGAAGAGAGATTCAAGTATTACGGGTTGCCAGTGGCTTTTATTATGAATGTATG 961
D 902 TTTCAGGAAGAGAGATTCAAGTATTACGGGTTGCCAGTGGCTTTTATTATGAATGTATG 961
QY 962 CATATGCTGGTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAACCTGAAAAA 1021
D 962 CATATGCTGGTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAACCTGAAAAA 1021
D 962 CATATGCTGGTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAACCTGAAAAA 1021
QY 1053 CATATGCTGGTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAACCTGAAAAA 1112
D 1053 CATATGCTGGTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAACCTGAAAAA 1112
D 1053 CATATGCTGGTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAACCTGAAAAA 1112
QY 1022 CCATTCCTCCCTTGGCAATATGATATCCATGGCCATTTGTCACCATTTGGCTATGTCGACAA 1081
D 1022 CCATTCCTCCCTTGGCAATATGATATCCATGGCCATTTGTCACCATTTGGCTATGTCGACAA 1081
D 1022 CCATTCCTCCCTTGGCAATATGATATCCATGGCCATTTGTCACCATTTGGCTATGTCGACAA 1081
QY 1113 CCATTCCTCCCTTGGCAATATGATATCCATGGCTATCATCATCAGTGGCTACGTACTGACAA 1172
D 1113 CCATTCCTCCCTTGGCAATATGATATCCATGGCTATCATCATCAGTGGCTACGTACTGACAA 1172
D 1113 CCATTCCTCCCTTGGCAATATGATATCCATGGCTATCATCATCAGTGGCTACGTACTGACAA 1172
QY 1082 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAG 1141
D 1082 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAG 1141
D 1082 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAG 1141
QY 1173 ACGTGGCCTATTTTACCACCATCAGTGGGAGGAGCTGTCGAGTCCAGCGCCCTGGCGG 1232
D 1173 ACGTGGCCTATTTTACCACCATCAGTGGGAGGAGCTGTCGAGTCCAGCGCCCTGGCGG 1232
D 1173 ACGTGGCCTATTTTACCACCATCAGTGGGAGGAGCTGTCGAGTCCAGCGCCCTGGCGG 1232
QY 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTGCCC 1201
D 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTGCCC 1201
D 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTGCCC 1201
QY 1233 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTGCCC 1292
D 1233 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTGCCC 1292
D 1233 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTGCCC 1292
QY 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTCTATGTTG 1261
D 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTCTATGTTG 1261
D 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTCTATGTTG 1261
QY 1293 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTCTACGTCG 1352
D 1293 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTCTACGTCG 1352
D 1293 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTCTACGTCG 1352
QY 1262 CGTCTCGAGAGGGTCAACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTC 1321
D 1262 CGTCTCGAGAGGGTCAACCTTCCAGAAATCCTCTCTCATGATTCATGTCGCAAGCACACTC 1321
D 1262 CGTCTCGAGAGGGTCAACCTTCCAGAAATCCTCTCTCATGATTCATGTCGCAAGCACACTC 1321
QY 1353 CATCTCGAGAGGGCACCTTCCGGAATCCTCTCTATGATTCATGTCGCAAGCACACTC 1412
D 1353 CATCTCGAGAGGGCACCTTCCGGAATCCTCTCTATGATTCATGTCGCAAGCACACTC 1412
D 1353 CATCTCGAGAGGGCACCTTCCGGAATCCTCTCTATGATTCATGTCGCAAGCACACTC 1412
QY 1322 CTCTACCACTGTTTATTTGTTTGCACCTTTTGACAAATGATATGCTCTCTCTGGAGACC 1381
D 1322 CTCTACCACTGTTTATTTGTTTGCACCTTTTGACAAATGATATGCTCTCTCTGGAGACC 1381
D 1322 CTCTACCACTGTTTATTTGTTTGCACCTTTTGACAAATGATATGCTCTCTCTGGAGACC 1381
QY 1413 CTCTGCCAGCTGTTATTTGTTTGCATCCTCTGACGATGGTGTGCTCTCTCTCGGAGACC 1472
D 1413 CTCTGCCAGCTGTTATTTGTTTGCATCCTCTGACGATGGTGTGCTCTCTCTCGGAGACC 1472
D 1413 CTCTGCCAGCTGTTATTTGTTTGCATCCTCTGACGATGGTGTGCTCTCTCTCGGAGACC 1472
QY 1382 TCGACAGCTTTTGAATTTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1441
D 1382 TCGACAGCTTTTGAATTTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1441
D 1382 TCGACAGCTTTTGAATTTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1441
QY 1473 TCTATAGCTTCTAAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1532
D 1473 TCTATAGCTTCTAAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1532
D 1473 TCTATAGCTTCTAAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1532
QY 1442 CTGGGCTGATTTATCTTCGATACAAATGCCCCAGATATGATCGTCTCTTCAAGGTGCCAC 1501
D 1442 CTGGGCTGATTTATCTTCGATACAAATGCCCCAGATATGATCGTCTCTTCAAGGTGCCAC 1501
D 1442 CTGGGCTGATTTATCTTCGATACAAATGCCCCAGATATGATCGTCTCTTCAAGGTGCCAC 1501
QY 1533 CAGGACTGATTTATCTTCGATACAAACGCCCCAGATATGATCGTCTCTTCAAGGTGCCCTC 1592
D 1533 CAGGACTGATTTATCTTCGATACAAACGCCCCAGATATGATCGTCTCTTCAAGGTGCCCTC 1592
D 1533 CAGGACTGATTTATCTTCGATACAAACGCCCCAGATATGATCGTCTCTTCAAGGTGCCCTC 1592
QY 1502 TGTTCATCCAGCTTTGTTTCTTCCATGATGCTCTTCTCATGTTGGCTTCCCTCTCTATT 1561
D 1502 TGTTCATCCAGCTTTGTTTCTTCCATGATGCTCTTCTCATGTTGGCTTCCCTCTCTATT 1561
D 1502 TGTTCATCCAGCTTTGTTTCTTCCATGATGCTCTTCTCATGTTGGCTTCCCTCTCTATT 1561
QY 1593 TCTTCATCCCGGCACTATTTTCTTCCATGCTGCTCTTCTCATGTTGGCTTCCCTCTCTATT 1652
D 1593 TCTTCATCCCGGCACTATTTTCTTCCATGCTGCTCTTCTCATGTTGGCTTCCCTCTCTATT 1652
D 1593 TCTTCATCCCGGCACTATTTTCTTCCATGCTGCTCTTCTCATGTTGGCTTCCCTCTCTATT 1652
QY 1562 CGGACCCCATTTAGTACAGGATTTGGCTTCCATGCTCTCTCATGTTGGCTTCCCTCTCTATT 1621
D 1562 CGGACCCCATTTAGTACAGGATTTGGCTTCCATGCTCTCTCATGTTGGCTTCCCTCTCTATT 1621
D 1562 CGGACCCCATTTAGTACAGGATTTGGCTTCCATGCTCTCTCATGTTGGCTTCCCTCTCTATT 1621
QY 1653 CGGACCCCATTTAGTACAGGATTTGGCTTCCATGCTCTCTCATGTTGGCTTCCCTCTCTATT 1712
D 1653 CGGACCCCATTTAGTACAGGATTTGGCTTCCATGCTCTCTCATGTTGGCTTCCCTCTCTATT 1712
D 1653 CGGACCCCATTTAGTACAGGATTTGGCTTCCATGCTCTCTCATGTTGGCTTCCCTCTCTATT 1712
QY 1622 ATCTCTTTTATATATGGGACAAAGAACCCAGGTTGTTTAGAATAATGTCAGAGAAAAATAA 1681
D 1622 ATCTCTTTTATATATGGGACAAAGAACCCAGGTTGTTTAGAATAATGTCAGAGAAAAATAA 1681
D 1622 ATCTCTTTTATATATGGGACAAAGAACCCAGGTTGTTTAGAATAATGTCAGAGAAAAATAA 1681
QY 1713 ATCTCTTCTATGTTGGGACAAAGAACCCAGGTTGTTTAGAATAATGTCAGAGAAAAATAA 1772
D 1713 ATCTCTTCTATGTTGGGACAAAGAACCCAGGTTGTTTAGAATAATGTCAGAGAAAAATAA 1772
D 1713 ATCTCTTCTATGTTGGGACAAAGAACCCAGGTTGTTTAGAATAATGTCAGAGAAAAATAA 1772
QY 1682 CCAGAACATTTACAAATAACTGGAAGTTGTACCAGAGAGAGATA--AGTTATGAACATA 1738
D 1682 CCAGAACATTTACAAATAACTGGAAGTTGTACCAGAGAGAGATA--AGTTATGAACATA 1738
D 1682 CCAGAACATTTACAAATAACTGGAAGTTGTACCAGAGAGAGATA--AGTTATGAACATA 1738
QY 1773 CCAGAACATTTACAGATTTACTAGAGTTGTACCAGAGAGACTCTAAAGAAATATGAACAT 1832
D 1773 CCAGAACATTTACAGATTTACTAGAGTTGTACCAGAGAGACTCTAAAGAAATATGAACAT 1832
D 1773 CCAGAACATTTACAGATTTACTAGAGTTGTACCAGAGAGACTCTAAAGAAATATGAACAT 1832
QY 1739 ATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAATAGGATTTTACTTTC 1798
D 1739 ATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAATAGGATTTTACTTTC 1798
D 1739 ATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAATAGGATTTTACTTTC 1798
QY 1833 AATGCATCAAAAGCTTGGCCATCTGCC--AGGATTGAGATACAAAATGGATTTT----- 1886
D 1833 AATGCATCAAAAGCTTGGCCATCTGCC--AGGATTGAGATACAAAATGGATTTT----- 1886
D 1833 AATGCATCAAAAGCTTGGCCATCTGCC--AGGATTGAGATACAAAATGGATTTT----- 1886
QY 1799 ATTTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGTATAAACAAGAGAGTCAAGTATTTT 1858
D 1799 ATTTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGTATAAACAAGAGAGTCAAGTATTTT 1858
D 1799 ATTTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGTATAAACAAGAGAGTCAAGTATTTT 1858
QY 1887 -----ATTCCAAGAAAACACAACTGATGATGAGCTAAAGAAATCAGTTATCTC 1936
D 1887 -----ATTCCAAGAAAACACAACTGATGATGAGCTAAAGAAATCAGTTATCTC 1936
D 1887 -----ATTCCAAGAAAACACAACTGATGATGAGCTAAAGAAATCAGTTATCTC 1936
QY 1859 TATTCATATATTTTACATATTTTGAACATTTTGAACATTTTGAAGAAATTTAGTATAACTCTATGT 1918
D 1859 TATTCATATATTTTACATATTTTGAACATTTTGAACATTTTGAAGAAATTTAGTATAACTCTATGT 1918
D 1859 TATTCATATATTTTACATATTTTGAACATTTTGAACATTTTGAAGAAATTTAGTATAACTCTATGT 1918
QY 1937 TATTCATATCTCTAGCGTATTTCAAAATTAATTTTGAAGCAACTTACTGTAACCTCAAGT 1996
D 1937 TATTCATATCTCTAGCGTATTTCAAAATTAATTTTGAAGCAACTTACTGTAACCTCAAGT 1996
D 1937 TATTCATATCTCTAGCGTATTTCAAAATTAATTTTGAAGCAACTTACTGTAACCTCAAGT 1996
QY 1919 AGTTATAGAAAGTGAATATGAGTT-----ATTCTATGAGTCGACAAATTTTGAAGTCTC 1973
D 1919 AGTTATAGAAAGTGAATATGAGTT-----ATTCTATGAGTCGACAAATTTTGAAGTCTC 1973
D 1919 AGTTATAGAAAGTGAATATGAGTT-----ATTCTATGAGTCGACAAATTTTGAAGTCTC 1973
QY 1997 ATTTGTAGCAAGCTAATATGCAAGTCAACAGTCAACAGTCAAGGCAAGCTCACAGTCTTGTAGTCTA 2056
D 1997 ATTTGTAGCAAGCTAATATGCAAGTCAACAGTCAACAGTCAAGGCAAGCTCACAGTCTTGTAGTCTA 2056
D 1997 ATTTGTAGCAAGCTAATATGCAAGTCAACAGTCAACAGTCAAGGCAAGCTCACAGTCTTGTAGTCTA 2056
```


Db 429 CTCCTAAGGCGGTGCTCCAGAACACGGGCAGCGTGGCATGTCTCTGACCATCTGGACGG 488

QY 482 TGTGTGGGGTCCCTGTACATATTGGAGCTTTGTCTTATGTCTGAATTTGGGAACAATAA 541

Db 489 TGTGTGGGGTCCCTGTACATATTGGAGCTTTGTCTTATGTCTGAATTTGGGAACAATAA 548

QY 542 AGAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACCAAGCTTTTG 601

Db 549 AGAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACCAAGCTTTTG 608

QY 602 TACGAGTCTGGGTGGAACCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661

Db 609 TACGAGTCTGGGTGGAACCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 668

QY 662 CATTTGACGCTACATTTCTGGAACCATTTTTTATTCAATGTGAATCCCTGAAGTTCGCA 721

Db 669 CATTTGACGCTACATTTCTGGAACCATTTTTTATTCAATGTGAATCCCTGAAGTTCGCA 728

QY 722 TCAAGCTCATACAGCTGTGGGCATAACTGTAGTGTGCTCCTAAATAGCATGAGTGTCA 781

Db 729 TCAAGCTCATACAGCTGTGGGCATAACTGTAGTGTGCTCCTAAATAGCATGAGTGTCA 788

QY 782 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTCGTATAA 841

Db 789 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTCGTATAA 848

QY 842 TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACAGCAGAACTTTTAAAGACGCCT 901

Db 849 TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACAGCAGAACTTTTAAAGACGCCT 908

QY 902 TTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATG 961

Db 909 TTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATG 968

QY 962 CATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAA 1021

Db 969 CATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAA 1028

QY 1022 CCATTCCCTTGCATATGTATATCCATGGCCATTGTCCACATTTGGCTATGTGCTGACAA 1081

Db 1029 CCATTCCCTTGCATATGTATATCCATGGCCATTGTCCACATTTGGCTATGTGCTGACAA 1088

QY 1082 ATGTGGCTTACTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCAG 1141

Db 1089 ATGTGGCTTACTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCAG 1148

QY 1142 TGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTCCGATCTTTGTTGCC 1201

Db 1149 TGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTCCGATCTTTGTTGCC 1208

QY 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGCTCCAGGTTATTCTATGTTG 1261

Db 1209 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGCTCCAGGTTATTCTATGTTG 1268

QY 1262 CGTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGCAAGCACACTC 1321

Db 1269 CGTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGCAAGCACACTC 1328

QY 1322 CTCTACCAAGCTTATTTGTTTTCACCTTTGACAAATGATAATGCTCTTCTCTGGAGACC 1381

Db 1329 CTCTACCAAGCTTATTTGTTTTCACCTTTGACAAATGATAATGCTCTTCTCTGGAGACC 1388

QY 1382 TCGACAGTCTTTGAATTTCCCTCAGTTTTCAGGTTGGCTTTTATTGGCTGGCAGTTG 1441

Db 1389 TCGACAGTCTTTGAATTTCCCTCAGTTTTCAGGTTGGCTTTTATTGGCTGGCAGTTG 1448

QY 1442 CTGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGCTCTTCAAGGTGCCAC 1501

Db 1449 CTGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGCTCTTCAAGGTGCCAC 1508

QY 1502 TGTTCATCCAGCTTTGTTTTCCTTCATGCTCTTTCATGTTGCTCTTCCCTCTATT 1561

Db 1509 TGTTCATCCAGCTTTGTTTTCCTTCATGCTCTTTCATGTTGCTCTTCCCTCTATT 1568

QY 1562 CGGACCCATTAGTACAGGATTGGCTTCGTTCATCACTCTGACTGGAGTCCCTGCGTATT 1621

Db 1569 CGGACCCATTAGTACAGGATTGGCTTCGTTCATCACTCTGACTGGAGTCCCTGCGTATT 1628

QY 1622 ATCTCTTTATTATATGGACAAAGAACCCAGGTGGTTTGTAGATAATGTGAGAGAAAATAA 1681

Db 1629 ATCTCTTTATTATATGGACAAAGAACCCAGGTGGTTTGTAGATAATGTGAGAGAAAATAA 1688

QY 1682 CCAGAACATTACAAAATAATACTGGAGTTGTACCAGAGAGATAAGTTATGAACATAATG 1741

Db 1689 CCAGAACATTACAAAATAATACTGGAGTTGTACCAGAGAGATAAGTTATGAACATAATG 1748

QY 1742 GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAAATAGGGATTTTACTTCAAT 1801

Db 1749 GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAAATAGGGATTTTACTTCAAT 1808

QY 1802 TTCTGAAAGTCTAGAGAAATTACAACCTTGGTGATAAAACAAA 1843

Db 1809 TTCTGAAAGTCTAGAGAAATTACAACCTTGGTGATAAAACAAA 1850

RESULT 8

AB040875

LOCUS

DEFINITION

AB040875

ACCESSION

VERSION

KEYWORDS

SOURCE

AB040875.1 GI:13516845

Homo sapiens adult cDNA to mRNA, clone_lib:brain cDNA library

clone:hxCT.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (sites)

AUTHORS

Kim,J.Y., Chairoungdua,A., Cha,S.H., Segawa,H., Matsuo,H., Kim,D.K., Endou,H. and Kanai,Y.

TITLE

Human cystine/glutamate exchanger: cDNA cloning and upregulation by oxidative stress in glioma cells

JOURNAL

REFERENCE

2 (bases 1 to 2000)

AUTHORS

Direct Submission

TITLE

Submitted (01-APR-2000) Yoshikatsu Kanai, Kyorin University School of Medicine, Department of Pharmacology and Toxicology; 6-20-2 Shinkawa, Mitaka, Tokyo 181-8611, Japan (E-mail:ykanai@kyorin-u.ac.jp, Tel:+81-422-47-5511(ex.3453), Fax:+81-422-79-1321)

FEATURES

Location/Qualifiers

1..2000

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="hxCT"

/clone_lib="brain cDNA library"

/dev_stage="adult"

1..2000

/gene="hxCT"

136..1620

/gene="hxCT"

/note="Na+-independent amino acid transporter transmembrane protein"

/codon_start=1

/product="cystine/glutamate exchanger"

/protein_id="BAB40574.1"

/db_xref="GI:13516846"

/translation="MVRKPVVSTISKGGYLOGNVNRLPSLGNKEPPGQEKVQLKRVK TLLRGVSIIGTIIGAGIFISPGVLQNTGSVMSLTWTVCVLSLFGALSVALGT TIKKSGHYTYILEVFGPLPAFVWVWVLLIIRPAATAVISLAFGRYILEPFFIQCEI PELAIKITAVGITVVMVLSMSVSWASRIQIFLTFCKLTAILIIIVPGVMQLIKQT QNFKDAFSGRDSITRLPLAFYGYMAYAGWFLNFVTEVENPEKTIPLAICISMAI VTIGYLTNVAYFTTINAEELLSNAVAVTFSERLLGNFSLAVPIFVALSCFGSMNGG VFAVSRLEYVASREGHLPEILSMIHVRKHTPLPAVIVLHPLTMILFSGDLDLSLNLFL

[illegible]

```

Db 1819 TTCTGAAAGTCTAGAGAAUTACAACCTTTGGTGATAAACAAAA 1860

RESULT 7
AB026891
LOCUS      AB026891
DEFINITION Homo sapiens mRNA for cystine/glutamate transporter, complete cds.
ACCESSION AB026891
VERSION    AB026891.1 GI:5668544
KEYWORDS   cystine/glutamate transporter.
SOURCE     Homo sapiens fibroblast cDNA to mRNA.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (sites)
REFERENCE  Sato,H., Tamba,M., Kuriyama-Matsumura,K., Okuno,S. and Bannai,S.
            Molecular cloning and expression of human xCT, the light chain of
            Antioxidants and Redox Signaling 2, 665-671 (2000)
            2 (bases 1 to 1861)
AUTHORS    Sato,H. and Bannai,S.
TITLE      Direct Submission
JOURNAL    Submitted (30-APR-1999) Hideyo Sato, University of Tsukuba,
            Institute of Basic Medical Sciences; Tennodai 1-1-1, Tsukuba,
            Ibaraki 305-8575, Japan (E-mail:hideyo-s@md.tsukuba.ac.jp,
            Tel:81-298-53-3282, Fax:81-298-53-3039)
FEATURES   Location/Qualifiers
            source          1..1861
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /cell_type="fibroblast"
            gene            1..1861
                        /gene="xCT"
                        236..1741
                        /gene="xCT"
            CDS              /codon_start=1
                        /product="cystine/glutamate transporter"
                        /protein_id="BAA82628.1"
                        /db_xref="GI:5668545"
                        /translation="MVRKPVVSTISKGYLQGNVNGRIPSLGNKEPPGQEKVOLKRKV
                        TLLRGVSIIGTIIGAGIFISPKGVLQNTGSVGMSLTIWTCVGLSLFGALSYAELGT
                        TIKKSGGHYTYILEVFGPLPAFVRWVELLIIRPAATAVISLAFGRVILEPFIQCEI
                        PELAIKLITAVGITVVMVLSMSVSWARIOIFLTFCCKLTAILIIIVPGVMOLIKQGT
                        QNEKDAFSGRDSITRLPLAFYGYMAYAGWFLNFVTEVENPEKTIPLAICISMAI
                        VTIGVLTNVAYFTTINAEELLSNAVAVTFSERLNPLSLAVIFVALSCFGSMNGG
                        VFVSRFLFYVASREGHLPEILSMIHVRKHTPLPAVILVHLPTMIMLFGSLDLSLLNFL
                        SFARWLFIGLAVAGLIYLRKCPDMHRPFKVPFLFIPALFSFTCLEMVALSLYSDFPST
                        GIGFVITLTGVPAYLYLFIWDKKPRWFRIMSEKITRILQIILEVVEEDKL"

BASE COUNT 485 a 402 c 421 g 553 t
ORIGIN

Query Match      68.8%; Score 1540.8; DB 9; Length 1861;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1606; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 182 GAAAGCCTGTTGTGTCCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAACGGGA 241
      |||||
Db 243 GAAAGCCTGTTGTGTCCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAACGGGA 302
      |||||

QY 242 GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCAGGACGCCTTTTCAGGAAGAGACG 301
      |||||
Db 303 GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCAGGACGCCTGGGCAGGACG 340
      |||||

QY 302 CCTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
      |||||
Db 341 -----GAGAAAGTGCAGCTGAAGAGGAAAGTCA 368
      |||||

QY 362 CTTTACTGAGGGAGTCTCCATTATCATTTGACCATCATTTGGAGCAGGAATCTTTCATCT 421
      |||||
Db 369 CTTTACTGAGGGAGTCTCCATTATCATTTGACCATCATTTGGAGCAGGAATCTTTCATCT 428
      |||||

QY 422 CTCCTAAGGGCGTGCTCCAGAACACGGSCAGCGTGGGCATGTCTCTGACCATCTGGACGG 481
      |||||

```


Db	1325	CTCTACAGCTGTTATTGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACC	1384
QY	1382	TCGACAGCTTTTGAATTTCCCTCAGTTTTCGCCAGGTGGCTTTTATTGGCTGGCAGTTG	1441
Db	1385	TCGACAGCTTTTGAATTTCCCTCAGTTTTCGCCAGGTGGCTTTTATTGGCTGGCAGTTG	1444
QY	1442	CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCCCTTTCAAGGTGCCAC	1501
Db	1445	CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCCCTTTCAAGGTGCCAC	1504
QY	1502	TGTTTCATCCAGCTTTGTTTTCCTTCACATGCCTCTTCATGGTTGCCCTTTCCCTCTATT	1561
Db	1505	TGTTTCATCCAGCTTTGTTTTCCTTCACATGCCTCTTCATGGTTGCCCTTTCCCTCTATT	1564
QY	1562	CGGACCCATTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCTATT	1621
Db	1565	CGGACCCATTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCTATT	1624
QY	1622	ATCTCTTATTATATGGGACAAAGAACCCAGGTGGTTTGAATAATGTGAGAGAAATAA	1681
Db	1625	ATCTCTTATTATATGGGACAAAGAACCCAGGTGGTTTGAATAATGTGAGAGAAATAA	1684
QY	1682	CCAGAACATTACAAATAATACTGGAAGTGTACCAGAGAAGATAAGTTATGAACTAATG	1741
Db	1685	CCAGAACATTACAAATAATACTGGAAGTGTACCAGAGAAGATAAGTTATGAACTAATG	1744
QY	1742	GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCATT	1801
Db	1745	GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCATT	1804
QY	1802	TTCTGAAAGCTAGAGAAATTACAACCTTTGGTGATAAACAAGGAGTCAGTTATTTTAT	1861
Db	1805	TTCTGAAAGCTAGAGAAATTACAACCTTTGGTGATAAACAAGGAGTCAGTTATTTTAT	1864
QY	1862	TCATATATTTTAGCATATTTCGAACCTAATTTCTAAGAAATTTAGTTATAACTCTATGAGT	1921
Db	1865	TCATATATTTTAGCATATTTCGAACCTAATTTCTAAGAAATTTAGTTATAACTCTATGAGT	1924
QY	1922	TATAGAAAGTGAATATGCAGTTATTCTATGAGTCGCACAAATCTTGAGTCTCTGATACCT	1981
Db	1925	TATAGAAAGTGAATATGCAGTTATTCTATGAGTCGCACAAATCTTGAGTCTCTGATACCT	1984
QY	1982	ACCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGTCATTCTCTACAAATA	2041
Db	1985	ACCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGTCATTCTCTACAAATA	2044
QY	2042	TGTTAGCACGGCAAGAACCTTCAAAFTGAAGACTGAGATTTTCTGTATATATGGGTTT	2101
Db	2045	TGTTAGCACGGCAAGAACCTTCAAAFTGAAGACTGAGATTTTCTGTATATATGGGTTT	2104
QY	2102	TGTAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAAGTGTTTTCAATTCTGA	2161
Db	2105	TGT-AAGATGGTTTACACACTACAGATGTCTATACTGTGAAAAGTGTTTTCAATTCTGA	2163
QY	2162	AAAAAGCATACATCATGATTATGGCAAGAGGAGAGAGGTA	2204
Db	2164	AAAAAGCATACATCATGATTATGGCAAGAGGAGAGAGGAA	2206
RESULT 6			
AF200708	1874 bp mRNA linear PRI 01-DEC-2000		
LOCUS	AF200708		
DEFINITION	Homo sapiens calcium channel blocker resistance protein CCBRL mRNA, complete cds.		
ACCESSION	AF200708		
VERSION	AF200708.1 GI:11493651		
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1874)		
AUTHORS	Conklin,D.S. and Beach,D.H.		

TITLE	CCBR1, novel CD98 light chain implicated in redox control and calcium signaling		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1874)		
AUTHORS	Conklin,D.S. and Beach,D.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-NOV-1999) Cold Spring Harbor Labs, 1 Bungtown Rd, Cold Spring Harbor, NY 11724, USA		
FEATURES	Location/Qualifiers		
source	1..1874		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/cell_line="fibrosarcoma HT1080"		
	246..1751		
	/note="X-c transporter; CD98 light chain"		
	/codon_start=1		
	/product="calcium channel blocker resistance protein CCBRL"		
	/protein_id="AAG35592.1"		
	/db_xref="GI:11493652"		
	/translation="MVRKPVYSTISKGYLQNVGRPLSLGNKEPPQEKVQLKRKV TLLRGVSIIGTIIGAGIFISPKVLQNTGSVMSLTITWTCVLSLFGALSVAELGT TIKKSGGHYTYILEVEGFLPAPVRVWVVELLIIRPAATAVISLAFRYILEPFIQCEI PELAIKLITAVGITVVMVLSMSVSWSARIQIFLTCKLTAILIIVPGVMQLIKQOT QNFKDAFSGRDSITRLPLAFYGYMYAGWFYLFNFTVEENPEKTIPLAICISMAI VTIGVLTNVAYFTTINAEELLSNAVAVTFSERLGNFSLAVPIFVALSCFGSMNGG VFAVSRFLFVASREGHLEILSMIHVKHTPLPAVIVLHPLTMIMFLSGDLSLNLFL SFARWLFIGLAVAGLIYLRKCPDMHRPKVPLFIPALFSFTCLFMVALSLYSDPFT GIGFVITLTGVPAYLYFIWDKKPRFRIMSEKITRLQIILEVVEEDKL"		
BASE COUNT	490 a	407 c	422 g 555 t
ORIGIN			
Query Match	68.9%; Score 1542.4; DB 9; Length 1874;		
Best Local Similarity	96.7%; Pred. No.: 0;		
Matches 1607; Conservative	0; Mismatches 1; Indels 54; Gaps 1;		
QY	182	GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGGAATGTTAACGGGA	241
Db	253	GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGGAATGTTAACGGGA	312
QY	242	GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGAGGAGCGCTTTTTCAGGAAGAGACG	301
Db	313	GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGAGGAGCGCTTTTTCAGGAAGAGACG	350
QY	302	CCTTTTCAGGAAGAGACGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA	361
Db	351	-----GAGAAAGTGCAGCTGAAGAGGAAAGTCA	378
QY	362	CTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATTATGGAGCAGGAATCTTCATCT	421
Db	379	CTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATTATGGAGCAGGAATCTTCATCT	438
QY	422	CTCCTAAGGGCGTCTCCAGAACACACGGGAGCGTGGGCATGTCTCTGACCATCTGGACGG	481
Db	439	CTCCTAAGGGCGTCTCCAGAACACACGGGAGCGTGGGCATGTCTCTGACCATCTGGACGG	498
QY	482	TGTGTGGGTCCTGTCTACTATTTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAA	541
Db	499	TGTGTGGGTCCTGTCTACTATTTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAA	558
QY	542	AGAAATCTGGAGGTCATTACACATATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTG	601
Db	559	AGAAATCTGGAGGTCATTACACATATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTG	618
QY	602	TACGAGTCTGGGTGGAACCTCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG	661
Db	619	TACGAGTCTGGGTGGAACCTCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG	678
QY	662	CATTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACCTTGCGA	721
Db	679	CATTGGACGCTACATCTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACCTTGCGA	738
QY	722	TCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGTGCTCTCTAAATAGCATGAGTGCA	781

```
Db 1854 TATAGAAAGTGAATATGCAGTTATCTATGATTCGCACAAATCTTGTGACTCTCTGATACCT 1913
QY 1982 ACCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTTGGTCAATCTCTACAAACATA 2041
Db 1914 ACCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTTGGTCAATCTCTACAAACATA 1973
QY 2042 TGTAGCACGGCAAGAACCTTCAAAATTGAAGACTGAGATTTTCTGTATATATGAGTTT 2101
Db 1974 TGTAGCACGGCAAGAACCTTCAAAATTGAAGACTGAGATTTTCTGTATATATGAGTTT 2033
QY 2102 TGTAAAGATGGTTTACACACTACACACTGCTATATGTTGAAAAGTGTCTTCAAAATTCTGA 2161
Db 2034 TGTAAAGATGGTTTACACACTATAGATGCTATATGTTGAAAAGTGTCTTCAAAATTCTGA 2093
QY 2162 AAAAAAGCATACATCATGATTATGGCAAAAGAGGAGAGAA 2200
Db 2094 AAAAAAGCATACATCATGATTATGGCAAAAGAGAGAGAA 2132

RESULT 5
AF252872 LOCUS AF252872 Homo sapiens cystine/glutamate transporter xct mRNA linear PRI 02-MAY-2001
DEFINITION Homo sapiens cystine/glutamate transporter xct mRNA, complete cds.
ACCESSION AF252872
VERSION AF252872.1 GI:13924719
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2482)
AUTHORS Chancy,C.D., Kekuda,R., Wang,H., Huang,W., Prasad,P.D., Smith,S.B.
and Ganapathy,V.
TITLE Structure, Function and Regulation of Human Cystine/Glutamate
Transporter in Retinal Pigment Epithelial Cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2482)
AUTHORS Wang,H., Prasad,P.D. and Ganapathy,V.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Biochemistry & Molecular Biology, Medical
College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
FEATURES
source
1..2482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
232..1737
/codon_start=1
/product="cystine/glutamate transporter xct"
/protein_id="AAK49111.1"
/db_xref="GI:13924720"
/translation="MVRKPVVSTISKGYLQGNVNRPLSLGNKEPPGQEKVQLRKV
TLLRGVSIITIGLIGAFIFSPKGVLQNTGSVMSLTITWTCVGLSLFGLSALSYAELGT
TIKSGGHTYILEVFGPLPAFVRVWVELLIIRPAATAVISLAFGRYILEPFFIOCEI
PELAIKLITAVGITVVMVLSMSVSWSARIQIFLTFCKLTAIIIIIVPGVMQLIKGQT
QNFKDAFSGRDSITRLPLAFYGYMAYAGWFYLNFTVEEVENPEKTIPLAICISMAI
VTIGYVLTVNAVYFTTINAEELLNSNAVAVTFSERLGNFSLAVPIFVALSCFGSMNGG
VFAVSRLFYVASREGHLPEILSMIHRKHTPLPAVIVLHPLTMIMLFSGDLDSLNLFL
SFARWLFIGLAVAGLIYLYRYKCPDMHRPFKVPFLIPALFSFTCLFMVALSLYSDPFST
GIGFVITLTGVPAYYLFIIWDKKPRFRIMSEKITRTLQIILEVVPEDKL"
BASE COUNT 735 a 484 c 521 g 742 t
ORIGIN

Query Match 84.3%; Score 1888.2; DB 9; Length 2482;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1965; Conservative 0; Mismatches 3; Indels 55; Gaps 2;

QY 182 GAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGGA 241
Db 239 GAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGGA 298
QY 242 GGTGCTTCCCTGGGCAACAAGGAGGCCACCTGGCAGGACGCCTTTTTCAGGAAGAGACG 301
```

```
Db 299 GGCTGCCTTCCCTGGGCAACAAGGAGGCCACCTGGGCAG----- 336
QY 302 CCTTTTCAGGAAGAGAGAGCGCCTTTTTCAGGAAGAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
Db 337 -----GAGAAAGTGCAGCTGAAGAGGAAAGTCA 364
QY 362 CTTTACTGAGGGAGTCTCCATTATCTGACACCATCATTTGGAGCAGGAATCTTTCATCT 421
Db 365 CTTTACTGAGGGAGTCTCCATTATCTGACACCATCATTTGGAGCAGGAATCTTTCATCT 424
QY 422 CTCCTAAGGGCGTCTCCAGAACACGGGCGACGCTGGGCATGTCTCTGACCATCTGGACGG 481
Db 425 CTCCTAAGGGCGTCTCCAGAACACGGGCGACGCTGGGCATGTCTCTGACCATCTGGACGG 484
QY 482 TGTGTGGGGTCCCTGTCACACTATTGAGGCTTTGTCTTATGCTGAATGGGAACAACATAAA 541
Db 485 TATGTGGGGTCCCTGTCACACTATTGAGGCTTTGTCTTATGCTGAATGGGAACAACATAAA 544
QY 542 AGAAATCTGGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTG 601
Db 545 AGAAATCTGGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTG 604
QY 602 TACGAGTCTGGGTGGAACCTCCTCATAATACGCCCTGCAGCTACTGTGTGATATCCCTGG 661
Db 605 TACGAGTCTGGGTGGAACCTCCTCATAATACGCCCTGCAGCTACTGTGTGATATCCCTGG 664
QY 662 CATTTGGACGCTACATCTCTGGAACCAATTTTATTCAATGTGAATCCCTGAACCTTGCGA 721
Db 665 CATTTGGACGCTACATCTCTGGAACCAATTTTATTCAATGTGAATCCCTGAACCTTGCGA 724
QY 722 TCAAGCTCATACAGCTGTGGGCATACACTAGTAGTGATGGTCTCTAAATAGCATGAGTGTC 781
Db 725 TCAAGCTCATACAGCTGTGGGCATACACTAGTAGTGATGGTCTCTAAATAGCATGAGTGTC 784
QY 782 GCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 841
Db 785 GCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 844
QY 842 TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCGAAGCTTTAAAGACGCCT 901
Db 845 TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCGAAGCTTTAAAGACGCCT 904
QY 902 TTTCAGGAAGAGATTCAGTATTACCGGTTGCCACTGGCTTTTATTATGGAATGATG 961
Db 905 TTTCAGGAAGAGATTCAGTATTACCGGTTGCCACTGGCTTTTATTATGGAATGATG 964
QY 962 CATATGCTGGCTGTTTACCTCAACTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAAA 1021
Db 965 CATATGCTGGCTGTTTACCTCAACTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAAA 1024
QY 1022 CCATTCCCTTGCATATATGTATATCCATGGCCATTGTCACCATTGGCTATGTGCTGACAA 1081
Db 1025 CCATTCCCTTGCATATATGTATATCCATGGCCATTGTCACCATTGGCTATGTGCTGACAA 1084
QY 1082 ATGTGGCTTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCAG 1141
Db 1085 ATGTGGCTTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTCAAATGCAGTGGCAG 1144
QY 1142 TGACCTTTTCTGAGCGGCTACTGGAAATTTCTCATTAGCAGTCCGATCTTTGTTGCC 1201
Db 1145 TGACCTTTTCTGAGCGGCTACTGGAAATTTCTCATTAGCAGTCCGATCTTTGTTGCC 1204
QY 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGCTCCAGGTTATTCTATGTTG 1261
Db 1205 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGCTCCAGGTTATTCTATGTTG 1264
QY 1262 CGTCTCGAGAGGGTCAACCTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACACTC 1321
Db 1265 CGTCTCGAGAGGGTCAACCTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACACTC 1324
QY 1322 CTCTACCAAGCTGTTATTGTTTGGACCCCTTTTGACAATGATAATGCTCTTCTCTGGAGACC 1381
```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20026 IMAGE:4562994"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
161..1666
/codon_start=1
/product="Similar to solute carrier family 7, (cationic
amino acid transporter, y+ system) member 11"
/protein_id="AAH12087.1"
/db_xref="GI:15082352"
/translation="MVRKPVVSTISKGYLQGVNGLPSLGNKEPPGQEKVQLKRV
TLRGVSIIGTIIGAGIFIPKGVQLQNTSGVMSLTITWVCGVLSLFGALSAYELGT
TIKSGGHYTYILEVFGPLFAFVRVWVELLIIRPAATAVISLAFGRYILEPFIQCEI
PELAKLITAVGITVVMVLSMSVSWSARIQIFLFCIKLTAILIIIVPGVMQIKGQT
QNFDAFSGRDSITRLPLAFYGYMYAYAGWFYLFNFVEEVENPEKTIPLAICISMAI
VTIGYVLTNVAYFTTINAEELLSNAVAVTFSERLGNFSLAVPIFVALSCFSGMNGG
VFAVSRFPYVASREGLPEILSMIHVRKHTPLPAVIVLHPLTMIMLFSGLDLSLNLFL
SFARWLFGLAVAGLIYLRKCPDMHRPFKVPFLFIPALFSFTCLFMVALSLYSDFEST
GIGFVITLTGVPAYYLFIIWKKPRFRIMSEKITRQLQIILEVVEEDKL"
BASE COUNT 599 a 443 c 452 g 661 t
ORIGIN
Query Match 84.3%; Score 1888.2; DB 9; Length 2155;
Best Local Similarity 96.9%; Pred. NO. 0;
Matches 1957; Conservative 0; Mismatches 8; Indels 54; Gaps 1;
QY 182 GAAAGCCTGTTGTGTCCACCATCTCCAAAGGAGGTACCTGCAGGGAAATGTTAACGGGA 241
Db |
QY 168 GAAAGCCTGTTGTGTCCACCATCTCCAAAGGAGGTACCTGCAGGGAAATGTTAACGGGA 227
Db |
QY 242 GGCTGCCTTCCCTGGGCAACAGGAGCCACCTGGSCAGGACGCCCTTTTCAGGAAGAGACG 301
Db |
QY 228 GGCTGCCTTCCCTGGGCAACAGGAGCCACCTGGSCAG----- 265
QY 302 CCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAAGTGCAGCTGAAGAGGAAAGTCA 361
Db |-----GAGAAAGTGCAGCTGAAGAGGAAAGTCA 293
QY 362 CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCACATCATTTGGAGCAGGAATCTTCATCT 421
Db |
QY 294 CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCACATCATTTGGAGCAGGAATCTTCATCT 353
QY 422 CTCCTAAGGGCGTGTCTCCAGAACACGGGCGAGCGTGGGCATGTCTCTGACCACCTCTGGACGG 481
Db |
QY 354 CTCCTAAGGGCGTGTCTCCAGAACACGGGCGAGCGTGGGCATGTCTCTGACCACCTCTGGACGG 413
QY 482 TGTGTGGGGTCTGTGTACATATTGGAGCTTTTGTCTTATGTCTGAATTGGGAACAACATAA 541
Db |
QY 414 TGTGTGGGGTCTGTGTACATATTGGAGCTTTTGTCTTATGTCTGAATTGGGAACAACATAA 473
QY 542 AGAAATCTGGAGGTCATTACACATATATTTTGGAAAGTCTTTTGGTCCATTACCACTTTTG 601
Db |
QY 474 AGAAATCTGGAGGTCATTACACATATATTTTGGAAAGTCTTTTGGTCCATTACCACTTTTG 533
QY 602 TACGAGTCTGGGTGGAACCTCCATATAATACGCCCTGCAGCTACTGTGTGATATCCCTGG 661
Db |
QY 534 TACGAGTCTGGGTGGAACCTCCATATAATACGCCCTGCAGCTACTGTGTGATATCCCTGG 593
QY 662 CATTTGGAGCGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACTTTCGA 721
Db |
QY 594 CATTTGGAGCGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACTTTCGA 653
QY 722 TCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGATGGTGCCTAAATAGCATGAGTGTC 781
Db |
QY 654 TCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGATGGTGCCTAAATAGCATGAGTGTC 713
QY 782 GCTGGAGCGCGCGGATCCAGATTTTCTTAACCTTTTGGAGCTCACAGCAATTTCTGATAA 841
Db |
QY 714 GCTGGAGCGCGCGGATCCAGATTTTCTTAACCTTTTTCAGAGCTCACAGCAATTTCTGATAA 773

QY 842 TTATAGTCCCTGGAGTTATGCAGCTAAATTAAGGTCAAACGCAGAACTTTAAAGACGCCT 901
Db |
QY 774 TTATAGTCCCTGGAGTTATGCAGCTAAATTAAGGTCAAACGCAGAACTTTAAAGACGCCT 833
Db |
QY 902 TTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGAATGTATG 961
Db |
QY 834 TTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGAATGTATG 893
QY 962 CATATGCTGGCTGGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA 1021
Db |
QY 894 CATATGCTGGCTGGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA 953
QY 1022 CCATTTCCCTTGGCAATATGATATATCCATGGCCATTGTCACCATTTGGCTATGTGCTGACAA 1081
Db |
QY 954 CCATTTCCCTTGGCAATATGATATATCCATGGCCATTGTCACCATTTGGCTATGTGCTGACAA 1013
QY 1082 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAG 1141
Db |
QY 1014 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAG 1073
QY 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1201
Db |
QY 1074 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1133
QY 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGCCTCTCCAGGTTATTTCTATGTTG 1261
Db |
QY 1134 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGCCTCTCCAGGTTATTTCTATGTTG 1193
QY 1262 CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTC 1321
Db |
QY 1194 CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTC 1253
QY 1322 CTCTACCAGCTGTTATTTGTTTGGACCCCTTTTGACAATGATGATGCTCTTCTCTGGAGACC 1381
Db |
QY 1254 CTCTACCAGCTGTTATTTGTTTGGACCCCTTTTGACAATGATGATGCTCTTCTCTGGAGACC 1313
QY 1382 TCGACAGCTCTTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1441
Db |
QY 1314 TCGACAGCTCTTTTGAATTTCCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTG 1373
QY 1442 CTGGGCTGATTTATCTTCGATACAAATGCCCCAGATATGCATCGTCTTCAAGGTGCCAC 1501
Db |
QY 1374 CTGGGCTGATTTATCTTCGATACAAATGCCCCAGATATGCATCGTCTTCAAGGTGCCAC 1433
QY 1502 TGTTCATCCAGCTTTGTTTTCCTTCACATGCCTCTTCATGGTGGCTTCCCTCTATT 1561
Db |
QY 1434 TGTTCATCCAGCTTTGTTTTCCTTCACATGCCTCTTCATGGTGGCTTCCCTCTATT 1493
QY 1562 CGGACCCATTTAGTACAGGATTGGCTTCGTCATGCCTCTCTACTGGAGTCCCTGCGTATT 1621
Db |
QY 1494 CGGACCCATTTAGTACAGGATTGGCTTCGTCATGCCTCTCTACTGGAGTCCCTGCGTATT 1553
QY 1622 ATCTCTTTATATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTGAGAGAAATAA 1681
Db |
QY 1554 ATCTCTTTATATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTGAGAGAAATAA 1613
QY 1682 CCAGAACATTTACAAATAATACTGGAAGTGTACCAGAAAGATAAGTATGAACTAATG 1741
Db |
QY 1614 CCAGAACATTTACAAATAATACTGGAAGTGTACCAGAAAGATAAGTATGAACTAATG 1673
QY 1742 GACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGATTTTACTTTCATT 1801
Db |
QY 1674 GACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGATTTTACTTTCATT 1733
QY 1802 TTCTGAAAGCTTAGAGAAATTACAACCTTTGGTGATAACAAAAAGGAGTCACTTATTTTAT 1861
Db |
QY 1734 TTCTGAAAGCTTAGAGAAATTACAACCTTTGGTGATAACAAAAAGGAGTCACTTATTTAT 1793
QY 1862 TCATATATTTTACATATTTGAACTAATTTCTTAAGAAATTTAGTTATAACTCTATGTAGT 1921
Db |
QY 1794 TCATATATTTTATCATATTTGAACTAATTTCTTAAGAAATTTAGTTATAACTCTATGTAGT 1853
QY 1922 TATAGAAAGTGAATATGCAAGTTATTTCTATGAGTGCACAAATTTCTTGAGTCTCTGATACCT 1981

QY 722 TCAAGCTCATACAGCTGTGGGATAAAGTGTAGTGTGCTCTAAATAGCATGAGTGTC 781
Db 494 TCAAGCTCATACAGCTGTGGGATAAAGTGTAGTGTGCTCTAAATAGCATGAGTGTC 553
QY 782 GCTGGAGCGCCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTCGTATA 841
Db 554 GCTGGAGCGCCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTCGTATA 613
QY 842 TTATAGCTCCCTGGAGTTATGAGCTAATTAAGGTCACAGCAACTTTAAAGACGCCT 901
Db 614 TTATAGCTCCCTGGAGTTATGAGCTAATTAAGGTCACAGCAACTTTAAAGACGCCT 673
QY 902 TTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATG 961
Db 674 TTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTATG 733
QY 962 CATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCTGAAAAA 1021
Db 734 CATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCTGAAAAA 793
QY 1022 CCATTCCTCCCTTGCAATATGTATATCCATGGCCATTTGTCAACCATTTGGCTATGTCTGACAA 1081
Db 794 CCATTCCTCCCTTGCAATATGTATATCCATGGCCATTTGTCAACCATTTGGCTATGTCTGACAA 853
QY 1082 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCACTGGCAG 1141
Db 854 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCACTGGCAG 913
QY 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTTCAATAGCATGCTCCGATCTTTGTTGCC 1201
Db 914 TGACCTTTTCTGAGCGGCTACTGGGAAATTTTCAATAGCATGCTCCGATCTTTGTTGCC 973
QY 1202 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTCATGTTG 1261
Db 974 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTCATGTTG 1033
QY 1262 CGTCTCGAGAGGCTCACTTCCAGAAATCCTCTCCATGATTCATGCTCCGCAAGCACACTC 1321
Db 1034 CGTCTCGAGAGGCTCACTTCCAGAAATCCTCTCCATGATTCATGCTCCGCAAGCACACTC 1093
QY 1322 CTCTACAGCTGTTTATGTTTTCACACCTTTTGACACCTTTGACAAATGATAATGCTCTCTGGAGACC 1381
Db 1094 CTCTACAGCTGTTTATGTTTTCACACCTTTTGACACCTTTGACAAATGATAATGCTCTCTGGAGACC 1153
QY 1382 TCGACAGCTCTTTTGAATTTCTCTCAGTTTTCGACAGGTTGGCTTTTATGCGGCTGGCAGTTG 1441
Db 1154 TCGACAGCTCTTTTGAATTTCTCTCAGTTTTCGACAGGTTGGCTTTTATGCGGCTGGCAGTTG 1213
QY 1442 CTGGGCTGATTTATCTCGATACAAATGCCAGATATGCATCGTCTCTCAAGGTGCCAC 1501
Db 1214 CTGGGCTGATTTATCTCGATACAAATGCCAGATATGCATCGTCTCTCAAGGTGCCAC 1273
QY 1502 TGTTTATCCAGCTTTGTTTTCCTTTCACATGCTCTTTCATGTTTGGCTTTCCCTCTATT 1561
Db 1274 TGTTTATCCAGCTTTGTTTTCCTTTCACATGCTCTTTCATGTTTGGCTTTCCCTCTATT 1333
QY 1562 CGGACCCATTTAGTACAGGATTTGGCTTCTGCTCATCACTCTGACTGGAGTCCCTGCGTATT 1621
Db 1334 CGGACCCATTTAGTACAGGATTTGGCTTCTGCTCATCACTCTGACTGGAGTCCCTGCGTATT 1393
QY 1622 ATCTCTTTATTTATGAGGACAAAGAACCCAGGTTGTTTAGAATAATGTACAGAGAAATAA 1681
Db 1394 ATCTCTTTATTTATGAGGACAAAGAACCCAGGTTGTTTAGAATAATGTACAGAGAAATAA 1453
QY 1682 CCAGAACATTTACAAATAATCTGGAAGTTGTACCAGAAAGATAAGTTATGAACATAATG 1741
Db 1454 CCAGAACATTTACAAATAATCTGGAAGTTGTACCAGAAAGATAAGTTATGAACATAATG 1513
QY 1742 GACTTGAGATCTTGGAATCTGCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAAT 1801
Db 1514 GACTTGAGATCTTGGAATCTGCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAAT 1573
QY 1802 TTCTGAAAGCTTAGAGAATTAACAATTTGGTGATAAACAACAAAGGAGTCAGTTATTTAT 1861

Db 1574 TTCTGAAAGCTTAGAGAATTAACAATTTGGTGATAAACAACAAAGGAGTCAGTTATTTAT 1633
QY 1862 TCATATATTTTAGCATATTCGAACTAATTTCTAAGAAATTTAGTTATAACTCTATGTAGT 1921
Db 1634 TCATATATTTTAGCATATTCGAACTAATTTCTAAGAAATTTAGTTATAACTCTATGTAGT 1693
QY 1922 TATAGAAAGTGAATATATGAGTTATTTCTATGAGTCGACAAATTTTGAGTCTCTGATACCT 1981
Db 1694 TATAGAAAGTGAATATATGAGTTATTTCTATGAGTCGACAAATTTTGAGTCTCTGATACCT 1753
QY 1982 ACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGTCATTTCTCTACACATA 2041
Db 1754 ACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGTCATTTCTCTACACATA 1813
QY 2042 TGTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 2101
Db 1814 TGTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 1873
QY 2102 TGTAAAGATGGTTTACACACTACAGATGCTATACTGTGAAAAGTGTGTTCAATCTGA 2161
Db 1874 TGTAAAGATGGTTTACACACTACAGATGCTATACTGTGAAAAGTGTGTTCAATCTGA 1933
QY 2162 AAAAAAGCATACATCATGATTATGCGAAAAGAGGAGAGAGTA 2204
Db 1934 AAAAAAGCATACATCATGATTATGCGAAAAGAGGAGAGAGTA 1976

BC012087 2155 bp mRNA linear PRI 06-AUG-2001
Homo sapiens, Similar to solute carrier family 7, (cationic amino acid transporter, y+ system) member 11, clone MGC:20026
IMAGE:4562994, mRNA, complete cds.

ACCESSION BC012087
VERSION BC012087.1 GI:15082351
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTp
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: g Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5668544.
Location/Qualifiers
1. .2155

FEATURES
source

Db 1327 CGACAGCTCTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTTTATTATGGGCTGGCAGTTGC 1386

QY 1443 TGGGCTGATTATCTTCGATACAAATGCCCCAGATATGCATCGTCCTTTTCAAGGTGCCACT 1502

Db 1387 TGGGCTGATTATCTTCGATACAAATGCCCCAGATATGCATCGTCCTTTTCAAGGTGCCACT 1446

QY 1503 GTTCATCCCAGCTTTGTTTTCTTCACATGCCCTCTTTCATGGTGGCCCTTTCCCTCTATTTC 1562

Db 1447 GTTCATCCCAGCTTTGTTTTCTTCACATGCCCTCTTTCATGGTGGCCCTTTCCCTCTATTTC 1506

QY 1563 GGACCCATTAGTACAGGGATTGGCTTCGTCTCATCACTCTGACTGGAGTCCCTGCGTATTA 1622

Db 1507 GGACCCATTAGTACAGGGATTGGCTTCGTCTCATCACTCTGACTGGAGTCCCTGCGTATTA 1566

QY 1623 TCTCTTTATTATATGGGACAAGAAACCCAGGTGGTTTAGAATAATGTACAGAGAAATAAC 1682

Db 1567 TCTCTTTATTATATGGGACAAGAAACCCAGGTGGTTTAGAATAATGTACAGAGAAATAAC 1626

QY 1683 CAGAACATTACAAATAATACTGGAAGTTGTACCAGAAGAGATAAGTTATGAACATAATGG 1742

Db 1627 CAGAACATTACAAATAATACTGGAAGTTGTACCAGAAGAGATAAGTTATGAACATAATGG 1686

QY 1743 ACTTGAGATCTGGCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCAATT 1802

Db 1687 ACTTGAGATCTGGCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCAATT 1746

QY 1803 TCTGAAAGTCTAGAGAATTACAACTTTGGTGATATAACAAAAGGAGTCAGTTATTTTATT 1862

Db 1747 TCTGAAAGTCTAGAGAATTACAACTTTGGTGATATAACAAAAGGAGTCAGTTATTTTATT 1806

QY 1863 CATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTTAGTATAACTCTATGTAGTT 1922

Db 1807 CATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTTAGTATAACTCTATGTAGTT 1866

QY 1923 ATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTCTGAGTCTCTGATACCTA 1982

Db 1867 ATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTCTGAGTCTCTGATACCTA 1926

QY 1983 CCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGTCAATTTCTCTACAACATAT 2042

Db 1927 CCTATTGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGTCAATTTCTCTACAACATAT 1986

QY 2043 GTTAGCAGCGGAAAGAACCTTCAAATTAAGACTGAGATTTTCTGTATATATGGGTTT 2102

Db 1987 GTTAGCAGCGGAAAGAACCTTCAAATTAAGACTGAGATTTTCTGTATATATGGGTTT 2046

QY 2103 GTAAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAGTGTCTTCAATTCGAA 2162

Db 2047 GTAAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAGTGTCTTCAATTCGAA 2106

QY 2163 AAAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAGGTA 2204

Db 2107 AAAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAGAA 2148

RESULT 3

HSA277882 3144 bp mRNA linear PRI 04-JAN-2002

LOCUS Homo sapiens mRNA for cystine/glutamate transporter (XCT gene).

DEFINITION AJ277882

ACCESSION AJ277882

VERSION AJ277882.1 GI:18073361

KEYWORDS cystine/glutamate transporter; XCT gene.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Borsani, G., Manzoni, M., Palacin, M., Pineda, M. and Gasol, E.

Unpublished

2 (bases 1 to 3144)

REFERENCE Borsani, G., Manzoni, M., Palacin, M., Pineda, M. and Gasol, E.

AUTHORS Bassi, M.T.

TITLE Direct Submission

JOURNAL Submitted (11-MAY-2000) Bassi M.T., Telethon Institute of Genetics

FEATURES

Location/Qualifiers

1..3144

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="4"

/map="4q28-q32"

/cell_line="NT2, undifferentiated teratocarcinoma cell line"

gene

1..3128

/gene="XCT"

CDS

1..1506

/gene="XCT"

/codon_start=1

/product="cystine/glutamate transporter"

/protein_id="CAC81905.1"

/db_xref="GI:18073362"

/translation="MVRKPVVSTISKGYLQNVNRLPSLGNKEPPGQEKVQLKRKV

TLLRGVSIIGTIIGAGIFISPKVLQNTGSVMSLTITWVCGVLSLFGALSVAELGT

TIKSGGHVYILEVFGPLPAFVRVWVLLIIRPAATAVISLAFGRYILEPFFIQCEI

PELAIKLITAVGITVVMVLSMSVSWSARIOIFLTKLTAILLIIVPGVMQLIKGQT

QNFDAFSGRDSSITRLPLAFYGYMYAYAGWFYLNFTVEEVENPEKTIPLAICISMAI

VTIGVLTNVAYFTTINAEELLSNAVAVTFSERLGNFSLAVPIFVALSCFSMNGG

VFAVSRFLFYVASREGHLPEILSMIHVRKHTPLPAVIVLHPLTMIMLFSGLDLSLNLFL

SFARWLFIGLAVAGLIYLRKCPDMHRPFKVPFLFIPALFSFTCLFMVALSLSDPEST

GIGFVITLTGVPAAYLYLFIWDDKPRFRIMSEKIITRLQIILEVVPEDKL"

repeat_region 2857..3127

/rpt_family="Sine/Alu"

polyA_site 3128

/gene="XCT"

BASE COUNT 912 a 602 c 649 g 981 t

ORIGIN

Query Match 84.9%; Score 1900.2; DB 9; Length 3144;

Best Local Similarity 97.2%; Pred. No. 0;

Matches 1966; Conservative 0; Mismatches 3; Indels 54; Gaps 1;

QY 182 GAAAGCCTGTTGTCCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGGA 241

Db 8 GAAAGCCTGTTGTCCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGGA 67

QY 242 GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGACG 301

Db 68 GGCTGCCTTCCCTGGGCAACAAGGAGCGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGACG 105

QY 302 CCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361

Db 106 -----GAGAAAGTGCAGCTGAAGAGGAAAGTCA 133

QY 362 CTTTACTGAGGGGAGTCTCCATTATCATTTGGACCATCATTTGGAGCAGGAATCTTCATCT 421

Db 134 CTTTACTGAGGGGAGTCTCCATTATCATTTGGACCATCATTTGGAGCAGGAATCTTCATCT 193

QY 422 CTCCTAAGGGCGTCTCCAGAACACGGGCGAGCTGGGCGATGTCTCTGACCATCTGGACGG 481

Db 194 CTCCTAAGGGCGTCTCCAGAACACGGGCGAGCTGGGCGATGTCTCTGACCATCTGGACGG 253

QY 482 TGTGTGGGGTCTGTCACTATTTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAA 541

Db 254 TGTGTGGGGTCTGTCACTATTTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAA 313

QY 542 AGAAATCTGGAGGTCTACACATATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTG 601

Db 314 AGAAATCTGGAGGTCTACACATATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTG 373

QY 602 TACGAGTCTGGGTGGAACCTCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661

Db 374 TACGAGTCTGGGTGGAACCTCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 433

QY 662 CATTTGGACGCTACATTTCTGGAACCATTTTTTATTAATGTAATCCCTGAACCTTGCGA 721

Db 434 CATTTGGACGCTACATTTCTGGAACCATTTTTTATTAATGTAATCCCTGAACCTTGCGA 493

QY 1981 TACCTATTGGGTTAGGAGAAAAGACTAGACAATTACTATGTGGTCATTCTCTACAACAT 2040
|||||
Db 1981 TACCTATTGGGTTAGGAGAAAAGACTAGACAATTACTATGTGGTCATTCTCTACAACAT 2040
|||||
QY 2041 ATGTTAGCACGGCAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGGTT 2100
|||||
Db 2041 ATGTTAGCACGGCAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGGTT 2100
|||||
QY 2101 TTGTTAAAGATGGTTTTACACACTACAGATGTCTATACTGTGAAAAGTGTCTTAAATTCTG 2160
|||||
Db 2101 TTGTTAAAGATGGTTTTACACACTACAGATGTCTATACTGTGAAAAGTGTCTTAAATTCTG 2160
|||||
QY 2161 AAAAAAGCATAACATCATGATTATGGCAAAGAGGAGAGAAAGGTAGAGCTGTCTTAAATT 2220
|||||
Db 2161 AAAAAAGCATAACATCATGATTATGGCAAAGAGGAGAGAAAGGTAGAGCTGTCTTAAATT 2220
|||||
QY 2221 TATTAAAAAATAAAAAA 2239
|||||
Db 2221 TATTAAAAAATAAAAAA 2239
|||||
RESULT 2
AX321910
LOCUS AX321910 5981 bp DNA linear PAT 15-DEC-2001
DEFINITION Sequence 441 from Patent WO017295.
ACCESSION AX321910
VERSION AX321910.1 GI:17906518
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
Mannion,J. and Kalos,M.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 017295-A 441 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..5981
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1858 a 1029 c 1098 g 1996 t
ORIGIN
Query Match 92.9%; Score 2080.8; DB 6; Length 5981;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 2; Indels 54; Gaps 1;
QY 3 AGGTTGAAGTGACGACAGATCATGCCAGCCTGGTGACAGTGAGACTCTGTCTCAAACAG 62
|||||
Db 1 AGGTTGAAGTGACGACAGATCATGCCAGCCTGGTGACAGTGAGACTCTGTCTCAAACAG 60
|||||
QY 63 AATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGGAAATTCAGGCCAATTGTGGCA 122
|||||
Db 61 AATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGGAAATTCAGGCCAATTGTGGCA 120
|||||
QY 123 TAGATTTATCATATCTGGATTTTGGATTTTGTGTTTCTCATFCACTGGATTTCAGG 182
|||||
Db 121 TAGATTTATCATATCTGGATTTTGGATTTTGTGTTTCTCATFCACTGGATTTCAGG 180
|||||
QY 183 AAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGGAG 242
|||||
Db 181 AAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGGAG 240
|||||
QY 243 GCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGACGC 302
|||||
Db 241 GCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGAG----- 277
|||||
QY 303 CTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCAC 362
|||||

Db 278 -----GAGAAAGTCAGCTGAAGAGGAAAGTCAC 306
|||||
QY 363 TTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATTCTGGAGCAGGAATCTTCATCTC 422
|||||
Db 307 TTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATTCTGGAGCAGGAATCTTCATCTC 366
|||||
QY 423 TCCTAAGGGCGTGTCTCCAGAACACGGGCGAGCGTGGGCAATCTGTGCTATGCTGAATTTGGGAACAACATATAAA 542
|||||
Db 367 TCCTAAGGGCGTGTCTCCAGAACACGGGCGAGCGTGGGCAATCTGTGCTATGCTGAATTTGGGAACAACATATAAA 486
|||||
QY 483 GTGTGGGTCCTGTCTCAGTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATATAAA 542
|||||
Db 427 GTGTGGGTCCTGTCTCAGTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATATAAA 486
|||||
QY 543 GAAATCTGGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTGT 602
|||||
Db 487 GAAATCTGGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTGT 546
|||||
QY 603 ACGAGTCTGGGTGGAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGC 662
|||||
Db 547 ACGAGTCTGGGTGGAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGC 606
|||||
QY 663 ATTTGACGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACCTGCGAT 722
|||||
Db 607 ATTTGACGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACCTGCGAT 666
|||||
QY 723 CAAGCTCATACAGCTGTGGGCATAAAGTGTAGTGTGCTGCTAAATAGCATGAGTGTGAG 782
|||||
Db 667 CAAGCTCATACAGCTGTGGGCATAAAGTGTAGTGTGCTGCTAAATAGCATGAGTGTGAG 726
|||||
QY 783 CTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAAT 842
|||||
Db 727 CTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAAT 786
|||||
QY 843 TATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAAGCTCAAAGCAGACCTTTAAAGACGCTT 902
|||||
Db 787 TATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAAGCTCAAAGCAGACCTTTAAAGACGCTT 846
|||||
QY 903 TTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTATGC 962
|||||
Db 847 TTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTATGC 906
|||||
QY 963 ATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAGTAGAAAAACCTGAAAAAAC 1022
|||||
Db 907 ATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAGTAGAAAAACCTGAAAAAAC 966
|||||
QY 1023 CATTCCCTTGCATATATGATATCCATGGCCATTGTCACCATTTGGCTATGTCGACAAA 1082
|||||
Db 967 CATTCCCTTGCATATATGATATCCATGGCCATTGTCACCATTTGGCTATGTCGACAAA 1026
|||||
QY 1083 TGTGGCCTACTTTACGACCATTAAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGT 1142
|||||
Db 1027 TGTGGCCTACTTTACGACCATTAAATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGT 1086
|||||
QY 1143 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCAITAGCAGTTCCGATCTTTGTTGCCCT 1202
|||||
Db 1087 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCAITAGCAGTTCCGATCTTTGTTGCCCT 1146
|||||
QY 1203 CTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTCCAGGTTATCTATGTTGC 1262
|||||
Db 1147 CTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTCCAGGTTATCTATGTTGC 1206
|||||
QY 1263 GTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTCC 1322
|||||
Db 1207 GTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTCC 1266
|||||
QY 1323 TCTACCAGCTGTTATGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCGAGACCT 1382
|||||
Db 1267 TCTACCAGCTGTTATGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCGAGACCT 1326
|||||
QY 1383 CGACAGCTTTTGAATTTCTCTCAGTTTTCAGGAGTGGCTTTTATTGGGCTGGCAGTTGC 1442
|||||

TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0172295-A 440 04-Oct-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .2239
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 619 a 444 c 493 g 683 t
ORIGIN
Query Match 100.0%; Score 2239; DB 6; Length 2239;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAGGTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAC 60
Db 1 GGAGGTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAC 60
QY 61 AGAATTAAAGGAAAAAGAAAGAAAAAGAGAGAGAGAGAAATTCAGGCCCAATTGTGG 120
Db 61 AGAATTAAAGGAAAAAGAAAGAAAAAGAGAGAGAGAGAAATTCAGGCCCAATTGTGG 120
QY 121 CATAGATTTTATCATATTTCTGGATTTTCTGGATTTTCTGGATTTTCTCATCTGATTC 180
Db 121 CATAGATTTTATCATATTTCTGGATTTTCTGGATTTTCTCATCTGATTC 180
QY 181 GGAAGCCTGTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGG 240
Db 181 GGAAGCCTGTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGGAAATGTTAACGGG 240
QY 241 AGGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTTCAGGAAGAGAC 300
Db 241 AGGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTTCAGGAAGAGAC 300
QY 301 GCCTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTC 360
Db 301 GCCTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTC 360
QY 361 ACTTTACTGAGGGGAGTCTCCATATCATTTGGCACCATCATTTGGCAGGAGAAATCTCATC 420
Db 361 ACTTTACTGAGGGGAGTCTCCATATCATTTGGCACCATCATTTGGCAGGAGAAATCTCATC 420
QY 421 TCTCCTAAGGGCGTGCTCCAGAACACGGGACGGTGGGCATGTCTCTGACCATCTGACG 480
Db 421 TCTCCTAAGGGCGTGCTCCAGAACACGGGACGGTGGGCATGTCTCTGACCATCTGACG 480
QY 481 GTGTGTGGGTCCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAATA 540
Db 481 GTGTGTGGGTCCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAATA 540
QY 541 AAGAAATCTGGAGGTCAATTACACATATATTTTGAAGTCTTTGGTCCATTACCACTTTT 600
Db 541 AAGAAATCTGGAGGTCAATTACACATATATTTTGAAGTCTTTGGTCCATTACCACTTTT 600
QY 601 GTACGAGTCTGGGTGGAACCTCCTCAATAATACGCCCTGCAGTACTGTGTGATATCCCTG 660
Db 601 GTACGAGTCTGGGTGGAACCTCCTCAATAATACGCCCTGCAGTACTGTGTGATATCCCTG 660
QY 661 GCATTTGGACGCTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCG 720
Db 661 GCATTTGGACGCTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTGCG 720
QY 721 ATCAAGCTCATACAGCTGTGGGCATTAAGTGTAGTGTGATGTCCTAAATAGCATGAGTGC 780
Db 721 ATCAAGCTCATACAGCTGTGGGCATTAAGTGTAGTGTGATGTCCTAAATAGCATGAGTGC 780
QY 781 AGCTGGAGCGCGGATCCAGATTTTCTTAACCTTTTTCAGAGTCCAGCAATCTCTGATA 840
Db 781 AGCTGGAGCGCGGATCCAGATTTTCTTAACCTTTTTCAGAGTCCAGCAATCTCTGATA 840
QY 841 ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGGCAGAACTTTAAAGACGCC 900
Db 841 ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGGCAGAACTTTAAAGACGCC 900

Db 841 ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGGCAGAACTTTAAAGACGCC 900
QY 901 TTTTCAGGAAGAGATTTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTAT 960
Db 901 TTTTCAGGAAGAGATTTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTAT 960
QY 961 GCATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCTGAAAAA 1020
Db 961 GCATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCTGAAAAA 1020
QY 1021 ACCATTCCCTTGCAATATGTATATCCATGGCCATTTGTACCATTTGGCTATGCTGACA 1080
Db 1021 ACCATTCCCTTGCAATATGTATATCCATGGCCATTTGTACCATTTGGCTATGCTGACA 1080
QY 1081 AATGTGGCTACTTTACGACCATTATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCA 1140
Db 1081 AATGTGGCTACTTTACGACCATTATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCA 1140
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTTGCC 1200
Db 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGGATCTTTGTTGCC 1200
QY 1201 CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGCTCCAGGTTATTTCTATGTT 1260
Db 1201 CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGCTCCAGGTTATTTCTATGTT 1260
QY 1261 GCGTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACT 1320
Db 1261 GCGTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACT 1320
QY 1321 CCTCTACGAGCTTATTGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGAC 1380
Db 1321 CCTCTACGAGCTTATTGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGAC 1380
QY 1381 CTCGACAGCTTTTGAATTTCTCAGTTTTCAGGTTGGCTTTTATTTGGCTGGCAGTT 1440
Db 1381 CTCGACAGCTTTTGAATTTCTCAGTTTTCAGGTTGGCTTTTATTTGGCTGGCAGTT 1440
QY 1441 GCTGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGTCTTTCAAGGTGCCA 1500
Db 1441 GCTGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGTCTTTCAAGGTGCCA 1500
QY 1501 CTGTTTATCCAGCTTTGTTTCTTCCATGCTCTTCTCATGGTTGCCCTTTTCCCTCTAT 1560
Db 1501 CTGTTTATCCAGCTTTGTTTCTTCCATGCTCTTCTCATGGTTGCCCTTTTCCCTCTAT 1560
QY 1561 TCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTAT 1620
Db 1561 TCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTAT 1620
QY 1621 TATCTCTTTATATATGGACAAGAAACCCAGGTGGTTTAGAATAATGTCAGAGAAAAA 1680
Db 1621 TATCTCTTTATATATGGACAAGAAACCCAGGTGGTTTAGAATAATGTCAGAGAAAAA 1680
QY 1681 ACCAGAACATTACAAATAACTGGAAGTTGTACCAGAAAGAGATAAGTATGAACATA 1740
Db 1681 ACCAGAACATTACAAATAACTGGAAGTTGTACCAGAAAGAGATAAGTATGAACATA 1740
QY 1741 GGACTTGAGATCTTGGCAATCTGCCCCAAGGGGAGACACAAAATAGGGATTTTACTTCAT 1800
Db 1741 GGACTTGAGATCTTGGCAATCTGCCCCAAGGGGAGACACAAAATAGGGATTTTACTTCAT 1800
QY 1801 TTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGATAAACAAGAGGTCAGTTATTTTA 1860
Db 1801 TTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGATAAACAAGAGGTCAGTTATTTTA 1860
QY 1861 TTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGATAAACAAGAGGTCAGTTATTTTA 1920
Db 1861 TTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGATAAACAAGAGGTCAGTTATTTTA 1920
QY 1921 TTATAGAAAGTGAATATGAGTATTTCTATGAGTCCGACAAATTTCTTGTGATCTCTGATACC 1980
Db 1921 TTATAGAAAGTGAATATGAGTATTTCTATGAGTCCGACAAATTTCTTGTGATCTCTGATACC 1980

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 03:54:58 ; Search time 6493 Seconds
(without alignments)
10035.605 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2239	100.0	2239	6	AX321909	AX321909 Sequence
2	2080.8	92.9	5981	6	AX321910	AX321910 Sequence
3	1900.2	84.9	3144	9	HSA277882	AJ277882 Homo sapi
4	1888.2	84.3	2155	9	BC012087	BC012087 Homo sapi
5	1888.2	84.3	2482	9	AF252872	AF252872 Homo sapi
6	1542.4	68.9	1874	9	AF200708	AF200708 Homo sapi
7	1540.8	68.8	1861	9	AB026891	AB026891 Homo sapi
8	1373.4	61.3	2000	9	AB040875	AB040875 Homo sapi
9	1208.6	54.0	2216	10	AB022345	AB022345 Mus muscu
10	529.2	23.6	64196	9	AC110804	AC110804 Homo sapi
11	379.2	16.9	3448	9	BC028216	BC028216 Homo sapi
12	379.2	16.9	6296	9	D87432	D87432 Human mRNA
13	367	16.4	2263	5	AF019906	AF019906 Xenopus l
14	366.4	16.4	1536	9	AB011263	AB011263 Homo sapi
15	366.4	16.4	2072	6	AR084930	AR084930 Sequence
16	366.4	16.4	2186	9	HSA018474	Y18474 Homo sapien
17	366.4	16.4	2237	9	HSA130718	AJ130718 Homo sapi
18	366.4	16.4	2245	9	AF092032	AF092032 Homo sapi
19	365.4	16.3	1609	9	AF077866	AF077866 Homo sapi
20	365.4	16.3	1621	9	AB017908	AB017908 Homo sapi
21	364.8	16.3	2098	9	BC003062	BC003062 Homo sapi
22	364.8	16.3	2105	9	BC010107	BC010107 Homo sapi
23	364.8	16.3	2447	9	AB020532	AB020532 Homo sapi
24	362.2	16.2	1524	6	E41390	E41390 Amino acid
25	362.2	16.2	1767	9	AB018542	AB018542 Homo sapi
26	362.2	16.2	2254	5	XLASUR4B	Y12716 Xenopus lae
27	362.2	16.2	4539	6	E41369	E41369 Amino acid
28	362.2	16.2	4539	9	AB018009	AB018009 Homo sapi
29	362.2	16.2	4670	9	AF104032	AF104032 Homo sapi
30	357.8	16.0	1553	10	AF200684	AF200684 Rattus no
31	357.8	16.0	2077	10	AB020520	AB020520 Rattus no
32	354.4	15.8	2045	9	AK025377	AK025377 Homo sapi
33	351.4	15.7	4039	4	AF174615	AF174615 Bos tauru
34	349	15.6	2144	10	MMU130943	AJ130943 Mus muscu
35	347.4	15.5	2168	10	BC014709	BC014709 Mus muscu
36	346.8	15.5	3423	10	AB023409	AB023409 Mus muscu
37	346.8	15.5	3441	10	AB017189	AB017189 Mus muscu
38	345.8	15.4	2158	10	MMU012754	AJ012754 Mus muscu
39	344.8	15.4	4117	6	E58469	E58469 Neutral ami
40	344.8	15.4	4117	10	AB024400	AB024400 Rattus no
41	343.8	15.4	3455	6	E31557	E31557 Neutral ami
42	343.8	15.4	3455	6	E41370	E41370 Amino acid
43	343.8	15.4	3455	10	AB015432	AB015432 Rattus no
44	343.8	15.4	3728	9	AF171669	AF171669 Homo sapi
45	343.8	15.4	4211	6	E58471	E58471 Neutral ami

ALIGNMENTS

RESULT 1
AX321909
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX321909
Sequence
AX321909
AX321909.1
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
Mannion,J. and Kalos,M.D.

AX321909
Sequence
AX321909
AX321909.1
GI:17906515
440 from Patent WO0172295.
2239 bp
DNA
Linear
PAT 15-DEC-2001

QY 340 GTGCAGCTGAAGAGGAAAGTCACTTTACTAGGGGAGTCTCCATTATCATTTGGCACCATC 399
Db 196 GTGACCCCTGCAGCGCAATATCACACTGATCAATGGTGTGGCCATCATAGTGGCACCATC 255
QY 400 ATTGGAGCAGGAATCTTCATCTCTCCTAAGGGCGTGTCCAGAACACGGGCAGCGTGGGC 459
Db 256 ATCGGTTTCGGGCATCTCGTGACGCCACCGCGGTGTCAAGGAAGCGGCTCGCCCGGA 315
QY 460 ATGTCTCTGACCATCTGGACGGTGTGTGGGTCCTGTCACTATTTGGAGCTTTGTCTTAT 519
Db 316 CTGTCTGTGTGGTGTGGGCTGTGTGGGCGTCTTCTCCATCGTGGGCGCACTGTGTCTAC 375
QY 520 GCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCAATTACACATATATTTTGGGAAGTC 579
Db 376 GCGGAGCTGGGCACATACCATCTCAAAGTCAGGCGGCGACTATGCCCTACATGCTAGAGGTC 435
QY 580 TTTGGTCCATTACCACTTTTGTACGAGTCTGGGTGGAACTCCTCATATAATACGCCCTGCA 639
Db 436 TACGGCTCGCTGCCCGCCTTCTCAAGCTCTGGATCGAGCTGCTCATCATTCGGCCCTCC 495
QY 640 GCTACTGCTGTGATATCCCTGGCATTTGGAGGCTACATTTCTGGAAACCATTTTATTCAA 699
Db 496 TCACAGTACATCGTGGCGCTGGTCTTCGCCACATACCTGCTCAAGCGGCTCTTCCCACT 555
QY 700 TGTGAATCCCTGAACCTTGGATCAAGCTCAATACAGCTGTGGGCATAAATGTAGTGATG 759
Db 556 TGTCCCGTCCCGAGGAGGCTGCCAAGCTCGTGGCCTGCCCTCTGGTGCTACTACTCAG 615
QY 760 GTCCTAAATAGCATGAGTGTACAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGC 819
Db 616 GCTGTGAACGTGTACAGTGTGAAGGCTGCTACCCGCTGTGCAGGATGCCCTTGGGGCTGCC 675
QY 820 AAGCTCACAGCAATCTGATAAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAA 879
Db 676 AAAGTGTGGCCCTGGCCCTCATCATCTGCTCGGCTTCATCCAGATGGAAAGGACATA 735
QY 880 -----ACGCAGAACTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAAGT 921
Db 736 GGACAAGGGGATGCATCCAACTGCACAGAAAGTTGTCTTTGAGGCACCAATCTGGAC 795
QY 922 ATTACGCGGTTGCCACTGGCTTTTATATGGAATGTATGCATATGCTGGCTGTTTAC 981
Db 796 GTGGGGAACATTTGTGGCATTTGTACAGTGGCCTCTTCGCCCTACGGAGGATGGAACAT 855
QY 982 CTCAACTTTGTACTGAAGAAGTAGAAAACCCCTGAAAAAACCAATCCCTTGGCAATATGT 1041
Db 856 CTGAATTTGTACGGAGGAGATGATCAACCCCTACAGGAACCTCCCTTGGCCATCATC 915
QY 1042 ATATCCATGGCCATTTGCACCAATTTGGCTATGTCTGACAAATGTGGCCCTACTTTACGACC 1101
Db 916 ATCTCCTTGCCCATTTGTACCCCTGGTCTATGTGTGACGAACCTGGGCTTGGGAACCTACCAC 975
QY 1102 ATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTA 1161
Db 976 CTGTCTACCAACCAAGATGCTGACATCTGAAGCCGTGGCTGTGGATTTTGGGAACCTACCAC 1035
QY 1162 CTGGGAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATG 1221
Db 1036 CTGGGAGTCAATCTCTGGATCATCTCCTGTCTTCGTGGGCTTGTCTGCTTGGCTCTGTC 1095
QY 1222 AACGGTGGTGTGTTGCTGTCTCCAGGTTATCTATGTTGCTCTCGAGAGGGTCACTT 1281
Db 1096 AATGGGTCTCTGTTCAAGTCTCCTCAAGACTGTTCTTCGTGGATCCAGGAGGGGCCACTG 1155
QY 1282 CCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTCTCTTACCAGCTGTTATGTT 1341
Db 1156 CCTTCCATCTCTCCATGATCCACCCACAGCTTCTTGACACCGGTGCCATCACTGGTGTTC 1215
QY 1342 TTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTC 1401
Db 1216 ACGTGTGTCATGACCCCTGATGTACGCCCTTCTCCAGAGACATCTTCTCCATCATCAACTTC 1275
QY 1402 CTCAGTTTTCACAGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTCGA 1461

Db 1276 TTCAGCTTCTTCAACTGGCTGTGTGTGGCCCTGGCCATCATCGGCATGATGGCTCCGC 1335
QY 1462 TACAAATGCCCCAGATATGCAATCGTCTCTTTCAAGGTGCCACTGTTCATCCCAGCTTTGTT 1521
Db 1336 TTTAAGAAAGCCTGAGCTGGAGCTCCCATCAAGGTGAATCTGGCCCTCCCAAGTCTCTTT 1395
QY 1522 TCCTTCACATGCCCTCTTCATGGTTCCTTTCCTCTATTCGGACCCATTTAGTACAGGG 1581
Db 1396 ATCCTGGCCTGCCCTCTTCTTCATCGCCGTGTCTTCTGGAAGACACCCCTGGAGTGGGC 1455
QY 1582 ATTGGCTTCGTTCATCACCTCTGACTGGAGTCCCTGCCGTATTCATCTCTTTATATATGGAC 1641
Db 1456 ATTGGCTTCGCCATCATCTCTCAGCGGGCTGCCCTGTCTACTTCTTTGGTGTGTGGTGA 1515
QY 1642 AAGAAACCCAGGTGGTT 1658
Db 1516 AACAAAGCCCAATGGAT 1532

Search completed: April 16, 2003, 06:28:27
Job time : 533 secs

QY	812	CCTTTTGAAGCTCACAGCAATTTCTGATAAATTAATAGTCCCTGGAGTTATGCAGCTAATTA	871
Db	602	CGCCGCGCAAGCUCUGGCCCGGCGCCUGAUCAUCCUGCGGGCUUGCUGCAGAUCCGGA	661
QY	872	AAGGTCAAACGCGCAGACTTTAAAGACGCGCTTTTTCAGGAAGAGATTCAAG-----TATTA	925
Db	562	AGGGUGAUGUGUCCAACUAGAUAUCCCAACUUCUAUUUGAAGGCACCAACUUGAUGUGG	721
QY	926	CGCGGTTGCCACTGGCTTTTATATATGGAATGATGTCATATGCTGGCTGGTTTACCTCA	985
Db	722	GGAACAUUGUCUGGCAUUAUACAGCGGCCUCUUUGCCUAUGGAGGAUGGAUAUACUUGA	781
QY	986	ACTTTGTTACTGAAGAAGTAGAAAACCCCTGAAAAAACCATTTCCCCTTGCAATATGTATAT	1045
Db	782	AUUUCGUCACAGGAGAAUGAUAACCCCUACAGAAACCCUGCCCCUGGCCAUCAUCAUCU	841
QY	1046	CCATGGCCATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACGATTA	1105
Db	842	CCCUGCCCAUCGUGACGUGGUGUACGUGCUGACCAACCCUGGCCUACUUCACCCUGU	901
QY	1106	ATGCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGG	1165
Db	902	CCACCGAGCAGUUGCUGGUCGCGGAGGCCGUGGCCGUGGACUUCGGGAACUACUACCCUGG	961
QY	1166	GAATTTCTCATTAGCAGTTCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCATGAACG	1225
Db	962	CGCUAUGUCCUGGAUCAUCCCGGUCUUCUGUGGGCGCUGUCCUGCGUCCGCUCAUUG	1021
QY	1226	GTGGTGTGTTGCTGCTCCAGGTTATTCTATGTTGGCTCTCCTGCTTTGGCTCCATGCCAG	1285
Db	1022	GGUCCUGUUCACAUCUCCAGGCUCUUCUUCUGUGGGGUCUCCGGGAAGGCCACCGGCCU	1081
QY	1286	AAATCCTCTCCATGATTCATGTCCGCAAGCACACTCCTCTACCAGCTGTTATTGTTTTC	1345
Db	1082	CCAUCCUCUCCAUGAUCACCCACAGCUCUCCACCCCGUGCCGUCUCCUGUUDUCAGU	1141
QY	1346	ACCCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCA	1405
Db	1142	GUGUGAUGACGCGUCUACGCCUUCUCCAAGGACAUCUUCUCCGUCAUCAACUUCUUA	1201
QY	1406	GTTTGTGCCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACA	1465
Db	1202	GCUUCUUAACUGGCGUCUGCGUGGCCCGUGGCCAUCAUCCGGAUGAUCUGGCGCGCCACA	1261
QY	1466	AATGCCCCAGATATGATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTCT	1525
Db	1262	GAAAGCCUGAGCUGAGCGGCCCAUCAAGGUGAACCCUGGCCCGCCUGUGUUCUUAUCC	1321
QY	1526	TCACATGCCTCTTCATGGTTGCCCTTTCCCTCTATTTCGGACCCCATTTAGTACAGGGATTG	1585
Db	1322	UGGCCUGCCUUCUGAUCGCCGUCUCCUUCUGGAAGACACCCGUGGAGUGUGGCAUCCG	1381
QY	1586	GCTTCGTCACTCTGACTGGAGTCCCTCGGTATTATCTCTTTATTATATGGGACAAGA	1645
Db	1382	GCUUCACCAUCAUCCUACAGCGGGCGUGCCCGUCUACUUCUUGGGGUGUGGUGGAAAAACA	1441
QY	1646	AACCCAGGTGGTT	1658
Db	1442	AGCCCAAGUGGCU	1454

RESULT 13
ABA09358
ID ABA09358 standard; cDNA; 4585 BP.
XX
AC ABA09358;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human 4F2 light chain homologue-encoding cDNA, SEQ ID NO:1134.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
haematopoiesis regulation; tissue growth; immunomodulator; activin;

inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
proliferation; metastasis; cancer; tumour; haematopoietic disorder;
myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
chronic inflammatory condition; proliferative retinopathy;
atherosclerosis; coronary heart disease; arterial ischaemia;
bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; antiinflammatory;
antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
antifungal; vulnerary; antiulcer; ss.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-457740/49.

P-PSDB; ABB12114.

Human proteins and DNA encoding sequences useful for preventing,
treating or ameliorating a medical condition in a mammalian subject
e.g. arthritis and cancer -

Claim 1; Page 918-919; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;
haematopoiesis regulatory activity; tissue growth activity;
immunomodulatory activity; activin- or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a cDNA encoding a
novel human polypeptide of the invention.

Db 672 CAAATGGGAACCCCTGGTACAAAGATATTTTACCTATGCTAAAGTATTGGCACTGATCGC 731
QY 840 AATTATAGTCCCTGGAGTTATGACGCTAATTAAGGTCAAACGCAGAACTTTTAAAGACGC 899
Db 732 GGTCACTGTCAGGCATTTGTAGACTTGGCCAGGAGCCTCTACTCATTTTGAGAATTC 791
QY 900 CTTTTTCAGGAAGAGATTCAAGTATTACGGCGTTTGGCCACTGGCTTTTATTATGGAATGTA 959
Db 792 CTTTGAGGGTTTCATCATTTTGCAGTGGGTGACATTTGCCCTGGCACTGTACTCAGCTCTGTT 851
QY 960 TGCATATGCTGGCTGGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAAAACCCCTGAAAA 1019
Db 852 CTCCTACTCAGGCTGGGACACCCCTCAACTATGTCACTGAAGAGATCAAGAATCCTGAGAG 911
QY 1020 AACCAATCCCTTGAATATGTAATATGTAATCCATGGCCATTGTACCACTATGGCTATGCTGAC 1079
Db 912 GAACCTGCCCTCTCCATTGGCATCTCCCATGCCATTGTGACCACTCATCTATATCTTGAC 971
QY 1080 AAATGTGGCTACTTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTTCAAAATGCACTGGC 1139
Db 972 CAATGTGGCTATATATCTGTGTAGACATGAGAGACATCTTGCCAGTGTGCTGTGC 1031
QY 1140 AGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTTGTGC 1199
Db 1032 TGTGACTTTTGCAGATCAGATATTTTGAATATTTAACTGGATAATTTCCACTGTCTAGTGC 1091
QY 1200 CCTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTTGTCTCTCCAGGTTATTTCTATGT 1259
Db 1092 ATTATCCTGTTTGGTGGCCTCAATGCTCCATTTGCGCTGCTTCTAGGCTTTTCTTTGT 1151
QY 1260 TGGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACAC 1319
Db 1152 GGGCTCAAGAGAAGGCCATCTCCCTGATGCCATCTGTCATGATGCCATTTGAGCGGTTCAAC 1211
QY 1320 TCCTCTACAGCTGTTATTTGTTTGCACCCCTTTTGACAAATGATAATGCTCTCTCTGAGAA 1379
Db 1212 ACCAGTGCTTCTCTGCTCTTCAATGGTATCATGGCATGTGATCTACTTGTGCTGGAAGA 1271
QY 1380 CCTCGACAGTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGT 1439
Db 1272 CATCTTCCAGCTCATTAACCTACTACAGCTTCACTGCTACTGTTCTTTTGTGGGCTTCTAT 1331
QY 1440 TGCTGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGGTGCC 1499
Db 1332 TGTGGGTCAGCTTTATCTGCGCTGGAGGAGCCTGATCGACCTCGTCCCTCAAGCTCAG 1391
QY 1500 ACTGTTCACTCCAGCTTTGTTTCTCTCATATGCCCTCTTCTATGGTTGCCCTTTCCCTCTA 1559
Db 1392 CGTTTCTCCGATTTGCTTCTGCTCTGCACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
QY 1560 TTCGGACCAATTTAGTACAGGATTTGGCTTCTGCTATCATCTGACTGGAGTCCCTGCGTA 1619
Db 1452 CAGTGATACATCAACTCCCTCATCGCATTTGCCATTGCTGCTCTCAGGCTGCTGCTGCTT 1511
QY 1620 TTATCTCTTTATTATA 1635
Db 1512 CTTCTCATCATCAGA 1527

RESULT 12
AAA08393
ID AAA08393 standard; RNA; 1524 BP.
XX
AC AAA08393;

XX 12-JUL-2000 (first entry)
DE Human L-type amino acid transporter 1 RNA sequence SEQ ID NO:26.
XX L-type amino acid transporter 1; LAT1; 4F2hc; 4F2 heavy chain;
KW neutral amino acid transporter; tumour; cancer; proliferation;
KW cell membrane surface 4F2 molecule; anticancer; ss.

XX OS Homo sapiens.
XX WO200014228-A1.
XX 16-MAR-2000.
XX 03-SEP-1999; 99WO-JP04789.
XX 03-SEP-1998; 98JP-0249993.
XX 02-SEP-1999; 99JP-0248546.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Endou H, Kanai Y;
XX WPI; 2000-256979/22.
XX P-PSDB; AAY82488.
XX Neutral amino acid transporter protein which conjugates with cell membrane protein 4F2 and operates independently of sodium ions is useful for screening potential cancer proliferation inhibitors -
XX Claim 13; Page 176-180; 189pp; Japanese.
XX The present invention describes a cell surface protein which is an L-type amino acid transporter-1 (LAT1), which mediates the transport of neutral amino acids, leucine, isoleucine, phenylalanine, methionine, tyrosine, tryptophan, valine and histidine, into the cell independently of sodium ions. The LAT1 protein conjugates with the cell membrane surface molecule 4F2. Antisense DNA or RNA, antibodies, antibody fragments, and peptide and non-peptide antagonists to LAT1 are useful as anticancer agents. The present sequence encodes human LAT1, which is specifically claimed in the present invention.
XX Sequence 1524 BP; 245 A; 516 C; 438 G; 325 U; 0 other;

Query Match 16.2%; Score 362.2; DB 21; Length 1524;
Best Local Similarity 38.0%; Pred. No. 4.8e-85;
Matches 506; Conservative 228; Mismatches 593; Indels 6; Gaps 1;
QY 332 GAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTG 391
Db 122 GCGAGGGGUGACCCUGCAGCGGAACAUCACGUGCUCAACGGGUGGCCAUCACUGG 181
QY 392 GCACCATCATTTGGAGCAGGAATCTTCACTCTCTCTAAGGGCGTGTCTCCAGAACACGGCA 451
Db 182 GGACCAUUAUCGGCUGGGCAUCUUCUGUGAGCGCCACGGCGUGUCUACAGGAGCGCU 241
QY 452 GCGTGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTCTACTATTGGAGCTT 511
Db 242 CGCCGGGCGUGGUGGUGGUGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
QY 512 TGTCTTATGCTGAATGGGAACAATAAAGAAATCTGGAGGTCTATACATATATTT 571
Db 302 UCUGCUACGGGAGCUGCGCACCAUCAUCUCCAAAUCCGGCGGCGGCGGCGGCGGCGGCGG 361
QY 572 TGGAAAGTCTTTGGTCCATTACCACTTTTGTACGAGTCTGGGTGGAACTCTCTATAATAC 631
Db 362 UGGAGGUCUACGGCUGCGCGCCGCUUCUACAGCUCUGGAGCUGGCGGCGGCGGCGGCGGCGG 421
QY 632 GCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 691
Db 422 GGCCUUAUCGCGAGUACAUUGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481
QY 692 TTATTCAATGTGAATCCCTGAACCTTGGATCGATCAAGCTCATACAGCTGTGGGCATAACTG 751
Db 482 UCCCCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
QY 752 TAGTGATGGTCTTAATAGCATGAGTGTGAGTGGAGCGCCCGGATCCAGATTTCTTAA 811
Db 542 UGCUCACGGCGGUGAACUGCUACAGCGUGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601

QY	900	CTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGTA	959
Db	792	CTTTGAGGGTTTCATCAATTTGCAGTGGGTGACATTTGCCCTGGCACTGTACTCAGCTCTGTT	851
QY	960	TGCATATGCTGGCTGGTTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCCCTGAAA	1019
Db	852	CTCCTACTCAGGCTGGGACACCCCTCAACTATGTCTACTGAAGAGATCAAGAATCCTGAGAG	911
QY	1020	AACCAATTCCTCCCTTGCATATATGTATATCCATGGCCATTTGTCACCAATTTGGCTATGTGCTGAC	1079
Db	912	GAACCTGCCCTCTCCATTTGGCATCTCCATGCCCATTTGTCACCATCATCTATATCTTGAC	971
QY	1080	AAATGTGGCCTACTTTACGACCAATTAATGCTGAGGAGTGTCTTTCAAATGCAGTGGC	1139
Db	972	CAATGTGGCCTATTATACTGTCTAGACATGAGAGACATCTTGGCCAGTGTGCTGTTC	1031
QY	1140	AGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGC	1199
Db	1032	TGTGACCTTTGCAGATCAGATATTTGGAATATTTAACTGGATAATTCACACTGTCACTTGC	1091
QY	1200	CCTCTCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATCTATGT	1259
Db	1092	ATTATCCTGTTTTGGTGGCCTCAATGCCTCCATTTGGGCTGCTTCTAGGCTTTTCTTTGT	1151
QY	1260	TGCGTCTCGAGAGGGTCACTTTCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACAC	1319
Db	1152	GGGCTCAAGAGAAGGCCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGGTTCA	1211
QY	1320	TCCTCTACCAAGCTGTATTGTTTTTGACCCCTTTGACAATGATAATGCTCTCTCTGGAGA	1379
Db	1212	ACCAGTGCCTTCTCTGCTTCAATGGTATCATGGCATTGATCTACTTGTGCGTGAAGA	1271
QY	1380	CCTCGACAGTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTATGGGCTGGCAGT	1439
Db	1272	CATCTTCCAGCTCATTAACACTACAGCTTTCAGTACTGGTTCCTTTGTGGGGCTTCTAT	1331
QY	1440	TGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCCCTTTCAAGGTGCC	1499
Db	1332	TGTGGGTCAAGCTTATCTGCGCTGGAAGGAGCGCTGATCGACCTCGTCCCCCTCAAGCTCAG	1391
QY	1500	ACTGTTTCATCCCAGCTTTGTTTTCTCTTCACATGGCTCTTCATGGTTGCCCTTTCCCTCTA	1559
Db	1392	CGTTTTCTTCCCGATTGTCTCTCTGCTCTGACCATCTTCCCTGGTGGCTGTCTCCACTTTA	1451
QY	1560	TTCGGACCCATTTAGTACAGGGATTGGCTTCGTTCATCACTCTGACTGGAGTCCCTGCGTA	1619
Db	1452	CAGTGATACTATCAACTCCCTCATCGGCATTGGCATTGCCCTCTCAGGCCTGCCCTTTTA	1511
QY	1620	TTATCTCTTTATTATA	1635
Db	1512	CTTCTCATCATCAGA	1527

RESULT 11

AXX86999
ID AAX86999 standard: cDNA: 2072 bp.

AX86999:

18-OCT-1999 (first entry)

DE Human amino acid permease homologue (AAPH) nucleotide sequence.

AA
KW Amino acid permease homologue; AAPH; cancer; inflammatory disorder;
KW autoimmune disorder; cell proliferation disorder; adenocarcinoma;
KW AIDS; Crohn's disease; rheumatoid arthritis; arteriosclerosis; human;
KW psoriasis; hepatitis; gene therapy; ss.

OS Homo sapiens.

xx	Key	Location/Qualifiers
FH		

FT	key	LOCATION/
142..1677	CDS	

FT
/*taq= a

FT	/product= "amino acid permease homologue (AAPH)"
XX	
PN	US5942399-A.
XX	
PD	24-AUG-1999.
XX	
PF	06-MAY-1998; 98US-0073362.
XX	
PR	06-MAY-1998; 98US-0073362.
XX	
PA	(INCY-) INCYTE PHARM INC.
XX	
PI	Corley NC, Hillman JL, Yue H;
XX	
DR	WPI; 1999-493508/41.
DR	P-PSDB; AAY27076.
XX	
PT	New human amino acid permease homolog (AAPH) polynucleotide and its
PT	complement useful for prevention and treatment of cancer, autoimmune
PT	and cell proliferation disorders
XX	
PS	Claim 4; Fig 1A-F; 33pp; English.

Query Match 16.4%; Score 366.4; DB 20; Length 2072;
Best Local Similarity 54.6%;
Pred. No. 4.4e-86;
Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;

QY	300	CGCCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGT	359
Db	192	CTCCCCCTTTGGGTGATGGGGCCAGCCCGGAGCAGGTGAAGCTGAAGAAGGAGAT	251
QY	360	CACCTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTTCAT	419
Db	252	CTCACTGCTTAACGGCGTGTGCCCTGATTGTGGGGAACATGATCGGCTCAGGCATCTTTGT	311
QY	420	CTCTCCTTAAGGGCGTGCTCCAGAAACACGGGCAGCGTGGGCATGTCTCTGACCATCTGGAC	479
Db	312	TTCCCCCAAGGGTGTGCTCATATACAGTGCCTCCTTTGGTCTCTCTCTGGTCACTGGGC	371
QY	480	GGTGTGTGGGTCTGTCTACTATTTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAT	539
Db	372	TGTCGGGGGCTCTTCTCCGTCTTTGGGGCCCTTTGTATTAGGGAACCTGGGCACCACCAT	431
QY	540	AAAGAAATCTGGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTT	599
Db	432	TAAGAAATCTGGGGCCAGCTATGCCTATATCCTGGAGGCCCTTTGGAGGATTCCTTGCTTT	491
QY	600	TGTACGAGTCTGGGTGGAACCTCCTCATATAACGCCCTGCAGCTACTGCTGTGATATCCCT	659
Db	492	CATCAGACTCTGGACCTCCCTGCTCATCATTTGAGCCCCACCAGCCAGGCCATCATTGGCAT	551
QY	660	GGCATTGTGGACGCTACATTCTGGAAACCATTTTATTCAATGTGAAATCCCTGAACTGTC	719
Db	552	CACCTTTGCCAACTACATGGTACAGCCTCTCTTCCCGAGCTGCTTCGCCCTTATGCTGTC	611
QY	720	GATCAAGCTCATTAACAGCTGTGGGCATAACTGTAGTATGGTCCCTAAATAGCATGAGTGT	779
Db	612	CAGCCGCCTGTGGCTGTCTGCTTAAACCTTCATTAACTGTGCTCATGT	671
QY	780	CAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATCTTGAT	839

Db 816 GAACCTGCCCCCTCTCCATTGGCATCTCCATGCCATGTGTACCATCATCTATATCTTGAC 875

QY 1080 AAATGTGGCCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGC 1139

Db 876 CAATGTGGCCTATTATACTGTCTAGACATGAGAGACATCTTGGCCAGTGATGCTGTTGC 935

QY 1140 AGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCTGATCTTTGTTGC 1199

Db 936 TGTGACTTTTGCAGATCAGATATTTGGAATATTTAACTGGATAATCCCACTGTCAGTTGC 995

QY 1200 CCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTATTCTATGT 1259

Db 996 ATTATCTGCTTTTGGTGGCCCTCAATGCCCTCCATTGTTGGCTGCTTCTAGGCTTTCTTTGT 1055

QY 1260 TGGCTCTCGAGAGGCTACCTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACAC 1319

Db 1056 GGGCTCAAGAGAAGGCCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGGTTTAC 1115

QY 1320 TCCCTTACCAGCTGTTATTGTTTGTGACCCCTTTTGACAATGATAATGCTCTCTCTGGAGA 1379

Db 1116 ACCAGTGCCTTCTCTGCTCTCAATGGTATCATGGCATGTGATCTACTTGTGCGTGAAGA 1175

QY 1380 CCTCGACAGTCTTTTGAATTTCTCCTCAGTTTTTGGCCAGTGGCTTTTATTTGGGCTGGCAGT 1439

Db 1176 CATCTCCAGCTCATTAACACTACTACAGCTTCAGCTACTGTTGCTTTTGTGGGCTTTCTAT 1235

QY 1440 TGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTTTCAAGGTGCC 1499

Db 1236 TGTGGTCAAGCTTTTATCTGCGTGGGAGGAGCCTGATCGACCTCGTCCCTCAAGCTCAG 1295

QY 1500 ACTGTTTCATCCAGCTTTGTTTCTCTTCCCTTCACATGCTTTCATGTTGCTTCCCTCTA 1559

Db 1296 CGTTTCTTCCGATTTGCTCTCTGCTCTGCACCATCTCTCTGTTGGTGTTCCTACTTA 1355

QY 1560 TTCGGACCCATTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTA 1619

Db 1356 CAGTGATACTATCAACTCCCTCATCGGCATTGCCATTGCCCTCTCAGGCTGCCCTTTTA 1415

QY 1620 TTATCTCTTTATTATA 1635

Db 1416 CTTCCTCATCATCAGA 1431

RESULT 10

AAZ23240

ID AAZ23240 standard; cDNA; 2072 BP.

XX

AC AAZ23240;

XX

DT 24-JAN-2000 (first entry)

XX

DE Human amino acid permease homolog (AAPH) nucleic acid sequence.

XX

KW Amino acid permease homolog; AAPH; cancer; inflammatory disorder; AIDS;

KW autoimmune disorder; anemia; asthma; atherosclerosis; diabetes mellitus;

KW multiple sclerosis; osteoarthritis; osteoporosis; rheumatoid arthritis;

KW ulcerative colitis; infection; cell proliferation disorder; human; ss.

XX

OS Homo sapiens.

XX

PN US5981242-A.

XX

PD 09-NOV-1999.

XX

PF 03-FEB-1999; 99US-0243920.

XX

PR 06-MAY-1998; 98US-0073362.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Yue H, Corley NC, Hillman JL;

XX

DR WPI; 1999-633325/54.

DR P-PSDB; AAY41159.

XX New human amino acid permease homolog, useful in the diagnosis, treatment and prevention of cancer, inflammatory/autoimmune disorders and cell proliferation disorders -

PT

PT and cell proliferation disorders -

XX

PS Disclosure; Fig 1A-F; 32pp; English.

XX

CC This cDNA encodes a human amino acid permease homolog (AAPH). The AAPH nucleic acid, protein, AAPH (ant)agonists are useful in the diagnosis, treatment and prevention of cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid or uterus; inflammatory/autoimmune disorders such as AIDS, Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, diabetes mellitus, Crohn's disease, atopic dermatitis, dermatomyositis, emphysema, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid arthritis, ulcerative colitis, viral, bacterial, fungal, parasitic, protozoal infections; and cell proliferation disorders such as actinic, keratosis, cirrhosis, arteriosclerosis, myelofibrosis, and primary thrombocythemia.

XX

SQ Sequence 2072 BP; 446 A; 537 C; 494 G; 595 T; 0 other;

Query Match 16.4%; Score 366.4; DB 20; Length 2072;

Best Local Similarity 54.6%; Pred. No. 4.4e-86;

Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;

QY 300 CGCCTTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGAAGAGGAAAGT 359

Db 192 CTCCTCCCTTTGGGTGATGGGCCAGCCAGGGCCGGAGCAGGTGAAGCTGAAGAAGGAGAT 251

QY 360 CACTTTACTGAGGGGAGTCTCCATTATCATTTGGACCACCATCATTTGGAGCAGGAATCTTCAT 419

Db 252 CTCACCTGCTTAACGGCGTGTGCCTGATTTGGGGGAACATGATCGGCTCAGGCATCTTTGT 311

QY 420 CTCTCCTAAGGGCGTGTCCAGAACACGGGCGAGCGTGGGCATGTCTCTGACCATCTGGAC 479

Db 312 TTCCCCCAAGGTGTCTCATATACAGTGCCTCTCTTTGGTCTCTCTCTGTCATCTGGGC 371

QY 480 GGTGTGTGGGGTCTCTCCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACAT 539

Db 372 TGTCGGGGGCTCTCTCCGTCTTTGGGGCCCTTTGTTATCGGGAAGTGGGCACCCCAT 431

QY 540 AAAGAAATCTGGAGGTCAATTACACATATATTTTGGAAAGCTTTTGGTCCATTACCAGCTTT 599

Db 432 TAAGAAATCTGGGCCAGCTATGCCTATATCTCTGGAGGCCCTTTGGAGGATTCCTTGTCTT 491

QY 600 TGTACGAGTCTGGGTGGAACCTCTCTATAATACGCCCTTGCAGCTACTGCTGTGATATCCCT 659

Db 492 CATCAGACTCTGGACCTCCCTGCTCATATTGAGCCCCACCCAGCCAGGCCATCATTTGCCAT 551

QY 660 GGCATTTGGACGCTACATTTCTGGAACCAATTTTATTAATGTGAATCCCTGAACCTTGC 719

Db 552 CACCTTTGCCAACTACATGTTACAGCTCTCTTCCGAGTGTCTTCCGCCCTTATGCTGC 611

QY 720 GATCAAGCTCATTAACAGCTGTGGGCATACTGATGATGTTCTCTAAATAGCATGAGTGT 779

Db 612 CAGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTTAACTTCACTTAAGTATGCTGCTATGT 671

QY 780 CAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTTCACAGCAATTTCTGAT 839

Db 672 CAAATGGGGAACCCCTGGTACAAGATATTTTCACTATGCTAAAGTATGGCACTGATCGC 731

QY 840 AATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCAAGACTTTAAAGACGC 899

Db 732 GGTATCGTTGCAGGCATTTAGACTTGGCCAGGGAGCCCTCTACTCATTTTGAGAATTC 791

Db 831 CAATGTGGCCTATTATACTGTGTAGACATGAGAGACATCTTGGCCAGTGATGCTGTTGC 890
QY 1140 AGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCCGATCTTTGTTGC 1199
Db 891 TGTGACTTTTGCAGATCAGATATTTGGAAATTTAACTGGATAATTCCTACTGTCAGTTGC 950
QY 1200 CCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATCTATGT 1259
Db 951 ATTATCCTGTTTGGTGGCCTCAATGCCCTCCATGTGGCTGCTTCTAGGCTTTTCTTTGT 1010
QY 1260 TGGCTCTCGAGAGGTCACCTTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCACAC 1319
Db 1011 GGGCTCAAGAGAGAGGCGCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGGTTTAC 1070
QY 1320 TCCTCTACAGCTGTTATGTTTGTGACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGA 1379
Db 1071 ACCAGTGCCTTCTCTGCTCTTCAATGGTATCATGGCATGATCTACTTGTGCGTGAAGA 1130
QY 1380 CCTCGACAGTCTTTTGAATTTCTCTCAGTTTGGCCAGGTGGCTTTTATTTGGGCTGGCAGT 1439
Db 1131 CATCTTCCAGCTCAATACTACTACAGCTTACAGCTACTGGTTCTTTTGGGGCTTTCTAT 1190
QY 1440 TGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTCTTCAAGGTGCC 1499
Db 1191 TGTGGTTCAGCTTTATCTGCGCTGGAAGGAGCCTGATCGACCTCGTCCCCTCAAGCTCAG 1250
QY 1500 ACTGTTTCATCCAGCTTTGTTTCTCTTCCATGCTCTTCTCATGGTTGCCCTTTCCCTCTA 1559
Db 1251 CGTTTCTTCCGATGTCTTCTGCTCTGCGCTCTGCGACCATCTTCTGTTGGTGTGTTTCA 1310
QY 1560 TTCGGACCCATTAGTACAGGGATTGGCTTCGTCATCATCTCTGATGGAGTCCCTGCGTA 1619
Db 1311 CAGTGATACTATCAACTCCCTCATCGGCATGCCATTGCCCTCTCAGGCTGCCCCCTTTTA 1370
QY 1620 TTATCTCTTTTATTATA 1635
Db 1371 CTTCCTCATCATCAGA 1386

RESULT 9
ID AAA11955 standard; cDNA; 1656 BP.
XX
AC AAA11955;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human monp-2 cDNA #2.
XX
KW Human; monp-2; cell proliferation; drug development; immune disease;
KW drug development; cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 46..1581
FT /*tag= a
FT /product= "monp-2"
XX
PN WO200017238-A1.
XX
PD 30-MAR-2000.
XX
PF 22-SEP-1999; 99WO-JP05205.
XX
PR 24-SEP-1998; 98JP-0270356.
XX
PA (TAIS) TAISHO PHARM CO LTD.
XX
PI Yoshimoto M, Takayama K;
XX
DR WPI; 2000-283547/24.

DR P-PSDB; AAY87630.
XX
PT Novel gene monp-2 encoding monocyte-specific protein with cell
PT proliferation promoting effect, useful in studying its expression and
PT immune function, and in developing drugs for e.g. immune diseases and
PT cancers -
XX
PS Example 1; Page 28-31; 34pp; Japanese.
XX
CC This invention describes a novel protein MONP-2 (I) and its encoding
CC gene monp-2 (II) with a cell proliferation promoting effect isolated
CC from a human peripheral blood. The new protein (I) and gene are useful
CC in studying the expression abnormality and dysfunction of (I) as well
CC as maintenance of immune function and development of drugs for e.g.
CC immune diseases and cancers. This sequence encodes the MONP-2 protein
CC described in the method of the invention.
XX
SQ Sequence 1656 BP; 335 A; 445 C; 400 G; 475 T; 1 other;

Query Match 16.4%; Score 366.4; DB 21; Length 1656;
Best Local Similarity 54.6%; Pred. No. 3.9e-86;
Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;
QY 300 CGCCTTTTTCAGGAAGAGACGCCCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGT 359
Db 96 CTCCCTTTTGGGTGATGGGGCCAGCCAGGGCCGGAGCAGGTGAAGCTGAAGAAGGAGAT 155
QY 360 CACTTTACTGAGGGGAGTCTCCATTATCATTTGGCAGCATCATTTGGAGCAGGAATCTTCAT 419
Db 156 CTCAGTGTAAACGGCGTGTGCTGCTGATTTGGGGAACATGATCGGCTCAGGCATCTTGT 215
QY 420 CTCTCTTAAGGGCGTGTCTCCAGAAACACGGGCGAGCTGGGCATGTCTCTGACCATCTGGAC 479
Db 216 TTCCCCCAAGGGTGTGCTCATATACAGTGCCTCCTTGGTCTCTCTCTGTCATCTGGC 275
QY 480 GGTGTGTGGGTCTCTGCTACTATTTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAAT 539
Db 276 TGTGCGGGGCTCTTCTCGCTCTTTGGGGCCCTTTGTTATGCGGAACCTGGSCACCACCAT 335
QY 540 AAAGAAATCTGGAGGTCAATACACATATATTTTGAAGTCTTTGGTCCATTTACCAGCTTT 599
Db 336 TAAGAAATCTGGGGCCAGCTATGCCTATATCTTGGAGGCCCTTTGGAGGATTCCTTGTCT 395
QY 600 TGTACGAGTCTGGGTGGAACCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCT 659
Db 396 CATCAGACTCTGGACCTCCCTGCTCATCATTTAGCCCAACAGCCAGGCCATCATTTGCCAT 455
QY 660 GGCATTTGGACGCTACATTTCTGGAACCATTTTATTTATTTCAATGTGAATCCCTGAATTC 719
Db 456 CACCTTTGCCAACTACATGTTACAGCTCTCTCTCCCGAGCTGCTTCGCCCTTATGCTGC 515
QY 720 GATCAAGCTCATTAACAGCTGTGGGCATAAATGATGATGGTCCCTAAATAGCATGAGTGT 779
Db 516 CAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
QY 780 CAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGAT 839
Db 576 CAAATGGGAACCCCTGGTACAGATATTTTCACTATGCTAAAGTATTGGCACTGATCGC 635
QY 840 AATTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCAACGAGCACTTTAAAGACGC 899
Db 636 GGTCACTGTTGCAGGCATTTAGACTTTGGCCAGGAGCCCTCTACTCATTTTGAATTC 695
QY 900 CTTTTCAGGAAGAGATTAAGTATTAACGGGTTGCCACTGGCTTTTATTTATGGAATGTA 959
Db 696 CTTTGAAGGTTTCATCATTTGAGTGGGTGACATTTGCCCTGGCAGTGTACTCAGCTCTGTT 755
QY 960 TGCATATGCTGGCTGTTTACCTCAACTTTTGTACTGAAGAAGTAGAAAAACCTGAAA 1019
Db 756 CTCCCTACTCAGGCTGGACACCCCTCAACTATGTCACCTGAAGAGATCAAGATCCTGAGAG 815
QY 1020 AACCATTTCCCTTGCATATGATATATCCATGGCCATTTGTCACCATTTGGCTATGCTGAC 1079
Db 1020 AACCATTTCCCTTGCATATGATATATCCATGGCCATTTGTCACCATTTGGCTATGCTGAC 1079

QY 1106 ATGCTGAGGAGCTGCTGCTTTCAAATGACAGTGGCAGTGACCTTTTCTGAGCGGCTACTGG 1165
Db 967 CCACCGAGCAGATGCTGCTCGTCCGAGGCGGTGGCTTCCGGAACATATCACCTGG 1026
QY 1166 GAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTTGGCTCCCATGAACG 1225
Db 1027 GCGTCATGCTCCTGGATCATCCCCGCTTCTCGTGGGCTGTCTGCTGCTCGCTCCGTCATG 1086
QY 1226 GTGGTGTGTTTGTCTCCAGGTTATTTCTATGTTGCTCTCGAGAGGCTCACCTTCCAG 1285
Db 1087 GGTCCCTGTTTACATCTCCAGGCTCTTCTTCTGGGGTCCCGGAAGCCACCTGCCCT 1146
QY 1286 AATCCTCTCCATGATTCATGTCGCAAGCACACTCTCTACCAAGTCTTATTTGTTGC 1345
Db 1147 CCATCCTCTCCATGATCCACCCACAGCTCCTCACCCCGTCCCTGCTGTTTCAAGT 1206
QY 1346 ACCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTCCTCA 1405
Db 1207 GTGTGATGACGCTGCTCTACGCTTCTTCCAAAGGACATCTTCTCCGTCATCAACTTCTTCA 1266
QY 1406 GTTTTGCCAGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACA 1465
Db 1267 GCTTCTCAACTGGCTCTGCTGGGCTTCAAGGTGCCACTGTTTCAATCCAGCTTTTCTTCT 1525
QY 1466 AATGCCACAGATATGATGCTGCTTCAAGGTGCCACTGTTTCAATCCAGCTTTTCTTCT 1525
Db 1327 GAAAGCCTGAGCTTGAGCGGCCCATCAAGGTGAACCTGGCCCTGCCTGTGTTCTTCTTCA 1386
QY 1526 TCACATGCTCTTTCATGTTGCTTCTTCCCTCTATTCCGACCCCATTTAGTACAGGGATTG 1585
Db 1387 TGGCCTGCTCTTCTGATGCGCTCTCTTCTTGGAGACACCCGTTGGAGTGTGGCATCG 1446
QY 1586 GCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTATCTCTTTATATATGGACAAGA 1645
Db 1447 GCTTCACCATCATCTCTCAGCGGCTGCCCTCTACTTCTTGGGGTCTGTTGGAAAAACA 1506
QY 1646 AACCCAGGTGGTT 1658
Db 1507 AGCCCAAGTGGCT 1519

RESULT 8
AA11954
ID AA11954 standard; cDNA; 1533 BP.
XX
AC AA11954;

XX 01-AUG-2000 (first entry)
DE Human monp-2 cDNA.
XX Human; monp-2; cell proliferation; drug development; immune disease;
KW drug development; cancer; ss.
XX Homo sapiens.
PN W0200017238-A1.
XX 30-MAR-2000.
XX 22-SEP-1999; 99WO-JP05205.
XX 24-SEP-1998; 98JP-0270356.
XX (TAIS) TAISHO PHARM CO LTD.
PI Yoshimoto M, Takayama K;
XX WPI; 2000-283547/24.
DR P-PSDB; AAY87630.
XX Novel gene monp-2 encoding monocyte-specific protein with cell

PT proliferation promoting effect, useful in studying its expression and
PT immune function, and in developing drugs for e.g. immune diseases and
PT cancers
XX Claim 3; Page 26-27; 34pp; Japanese.
XX This invention describes a novel protein MONP-2 (I) and its encoding
CC gene monp-2 (II) with a cell proliferation promoting effect isolated
CC from a human peripheral blood. The new protein (I) and gene are useful
CC in studying the expression abnormality and dysfunction of (I) as well
CC as maintenance of immune function and development of drugs for e.g.
CC immune diseases and cancers. This sequence encodes the MONP-2 protein
CC described in the method of the invention.
XX Sequence 1533 BP; 301 A; 420 C; 363 G; 448 T; 1 other;
SQ
Query Match 16.4%; Score 366.4; DB 21; Length 1533;
Best Local Similarity 54.6%; Pred. No. 3.8e-86;
Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;
QY 300 CGCCTTTTTCAGGAAGAGACGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGT 359
Db 51 CTCCTCTTGGGTGATGGGGCCAGCCAGGCGCGGAGCAGGTGAAGAGGAGAT 110
QY 360 CACTTTTACTGAGGGGAGTCTCCATTTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCAT 419
Db 111 CTCACCTGCTTAACGGCGTGTGCTGATTTGGGGAACATGATCGGCTCAGGCATCTTTGT 170
QY 420 CTCCTCTAAGGGCGTGTCCAGAAACAGGGCAGCGTGGGATGCTCTGACCATCTGGAC 479
Db 171 TTCCCCCAAGGGTGTGCTCATATACAGTGCCTCTTGTGCTCTCTCTGTCATCTGGC 230
QY 480 GGTGTGTGGGTCTCTCACTATTGGAGCTTTGTCTTATGCTGAATTTGGAACTAT 539
Db 231 TGTCGGGGGCTCTTCTCCGCTTTTGGGCGCTTTTGTATGCGGAACCTGGGCAACCAT 290
QY 540 AAAGAAATCTGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTT 599
Db 291 TAAGAAATCTGGGCGAGCTATGCTTATATCTTGGAGGCTTTGGAGGATTCCTTGTCTT 350
QY 600 GTACGAGTCTGGGTGGAACCTCTCATATAACGCCCTGACGCTACTGCTGTGATATCCCT 659
Db 351 CATCAGACTCTGGACCTCCCTGCTCATCATCTGAGCCACCCAGCCAGCCCATCATGCCC 410
QY 660 GGCATTTGGACGCTACATTTCTGGAACCACTTTTATTAATCTGAATCTCCCTGAACCTGC 719
Db 411 CACCTTTGCCAATACATGGTACAGCTCTCTTCCCGAGCTGCTTCCGCTTATGCTGC 470
QY 720 GATCAAGCTCATTACAGCTGTGGGCATTAAGTGTAGTGTGCTTAAATAGCATGAGTGT 779
Db 471 CAGCCGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
QY 780 CAGCTGGAGCGCCGATCCAGATTTTCTTAACCTTTTGAAGCTCAGCAATCTCTGAT 839
Db 531 CAAATGGGAACCCCTGTAAGATATTTTCACTATGCTAAAGTATTTGGCACTGATCGC 590
QY 840 AATTATAGTCCCTGGAGTTATGACGCTAATTAAGTCAACCGCAGAACTTTAAAGACGC 899
Db 591 GGTCACTGTTGCAGGCAATTTAGACTTGGCCAGGAGGCTCTACTCATTTTGAGAATTC 650
QY 900 CTTTTCAGGAAGAGATTAAGTATTACGCGGTGGCAGCTGGCTTTTATATGGAATGTA 959
Db 651 CTTTGGAGGTTTCACTATTTGAGTGGGTGACATTTGCCCTGGCACTGTACTCAGCTCTGTT 710
QY 960 TGCATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAGTAGAAACCCCTGAAAA 1019
Db 711 CTCCTACTCAGGCTGGGACACCCCTCAACTATGTCTACTGAAGAGATCAAGATCTCTGAG 770
QY 1020 AACCATTCCTTGGCAATATGATATCCATGGCCATTTGTCACCATTTGGCTATGCTGAC 1079
Db 771 GAACCTGCCCTCTCCATTTGGCATCTCCATGCCCATTTGTCACCATCATCTATCTTGAC 830
QY 1080 AATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTTCAATGCAGTGGC 1139

Db 316 AATGCTGAGGAGCTGCTGNTTTCAAATGCANNTGGCAGTGACCTTTTCTGAGCGGCTACT 375
QY 1164 GGGAAATTTCTCATAGCAGTTCCGATCTTTGTTGCCCTCTCCTGCTTGGCTCCATGAA 1223
Db 376 GGGAAATTTCTCATAGCAGATCCGATCTTTGTTGCCCTNTCCTGCTGGGCTCCATNAA 435
QY 1224 CGGTGGTGTGT-TTGCTGTCTCCAGGTTATCTATGTTGC--GTCTCGAGAGGGTCACT 1280
Db 436 CNGGGGTGTGTGCTGCTGCTCCAGGTTATCTATGTTGCCCTNTCCTGCTGGGCTCCATNAA 495
QY 1281 TCCAGAAATCCTCTCCATGATTTCATGTCGCGAAGCACACTCTCTPACCAAGCTGTTATTGT 1340
Db 496 TCCANAAATNCTCTCCATGATTTCATGTCGCGAAGCACACTNCTCTAC---ANTGGTNTTG 552
QY 1341 TTTGCACCCCTTGACAAATGATAATGCTCTCTCT-TGAGACCTCGACAGTCTTTTGAATT 1399
Db 553 TTTGCACCCCTTGACAAATGATAATGCTCTCTNTTTGGGAGACCTCGACAGTCTTTTNAATT 612
QY 1400 TCCTCAGTTTGGCAGGTGGCTTTTATTTGTTGGCTGGCAGTTGCTGGGC-TGATTATCTT 1458
Db 613 TACTCAAGGTTGGCAGGTGGCTTTTATTTGTTGGCTGGCAATTGCTGGGCTTGATTATCTT 672
QY 1459 CGATACAAATGCCAGATATGCATCG--TCCTTTCAAGGTGCACTGTTTCATCCCCAGCT 1515
Db 673 NGATNCAATGCCNANATATGCATCGTCCCTTTTCAAAGGTGCCCCCTGTTTCATCCCCACTT 732
QY 1516 TTGTTTT 1522
Db 733 TTNTTTT 739

RESULT 7
AAA08372

ID AAA08372 standard; cDNA; 4539 BP.

XX
AC AAA08372;

XX 12-JUL-2000 (first entry)

XX Human L-type amino acid transporter 1 nucleotide sequence SEQ ID NO:1.

DE L-type amino acid transporter 1; LAT1; 4F2hc; 4F2 heavy chain;

KW neutral amino acid transporter; tumour; cancer; proliferation;

KW cell membrane surface 4F2 molecule; anticancer; ss.

XX Homo sapiens.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT Neutral amino acid transporter protein which conjugates with cell
PT membrane protein 4F2 and operates independently of sodium ions is
PT useful for screening potential cancer proliferation inhibitors -
XX Claim 9; Page 125-132; 189pp; Japanese.
CC The present invention describes a cell surface protein which is an
CC L-type amino acid transporter-1 (LAT1), which mediates the transport
CC of neutral amino acids, leucine, isoleucine, phenylalanine, methionine,
CC tyrosine, tryptophan, valine and histidine, into the cell independently
CC of sodium ions. The LAT1 protein conjugates with the cell membrane
CC surface molecule 4F2. Antisense DNA or RNA, antibodies, antibody
CC fragments, and peptide and non-peptide antagonists to LAT1 are useful
CC as anticancer agents. The present sequence encodes human LAT1, which is
CC specifically claimed in the present invention.
XX SQ Sequence 4539 BP; 765 A; 1450 C; 1297 G; 998 T; 29 other;

Query Match 16.6%; Score 370.8; DB 21; Length 4539;
Best Local Similarity 54.4%; Pred. No. 4.5e-87;
Matches 725; Conservative 13; Mismatches 589; Indels 6; Gaps 1;

QY 332 GAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTG 391
Db 187 GCGAGGGCGTGACCTGCAGCGGAACATCACGCTGCTCAACGGCGTGGCCATCATCTGG 246
QY 392 GCACCATCATTTGGAGCAGGAATCTTCTCATCTCTCTTAAGGGCGTCTCCAGAACACGGCA 451
Db 247 GGACCATTATCGGCTCGGCATCTTCTGAGCGCCACGGCGTGTCTCAAGGAGCAGGCT 306
QY 452 GCGTGGGCGATGCTCTGACCATCTGGACGGTGTGTGGGTCTCTGACACTATTTGGAGCTT 511
Db 307 GCGCGGGCGTGGCGCTGGTGTGGCGCGCTGCGGCGCTCTTCSNATHGTNGGCGCGC 366
QY 512 TGTCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATACACATATATTT 571
Db 367 TCTGCTACGGAGCTCGGCACCACTCTCCAAATWSNGNGGCGGCTACGCTACATGC 426
QY 572 TGAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATAATAC 631
Db 427 TGGAGGTCTACGGCTCGTGCNCGNTTYCTCAAGCTCTGGATCGAGCTGCTCATCATCC 486
QY 632 GCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 691
Db 487 GGCCTTCAWSNCARTAYATCGTGGCCCTGGTCTTCCGCACTTCTGCTCAAGCCGCTNT 546
QY 692 TTATTCAATGTGAAATCCCTGAACCTTGGATCAAGCTCATTAAGCTGTGGGCATACTG 751
Db 547 TYCCNACCTGCCCGGTGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606
QY 752 TAGTGATGCTCTAAATAGCATGAGTGTGAGCTGGAGCGCGCGGATCCAGATTTTCTTAA 811
Db 607 TGCTCAGCGCGGTGAAGTGTCTACAGCTGAAGGCGGAGGAGGAGGAGGAGGAGGAGG 666
QY 812 CCTTTTCAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGAGCTAATTA 871
Db 667 CCGCGCGCAAGCTCCTGGCCCTGGCCCTGATCATCTCTGCTGGGCTTCTGTCAGATCGGA 726
QY 872 AAGTCAACGCGAGAACTTTAAAGAGCGCTTTTTCAGGAAGAGATTCAGG-----TATTA 925
Db 727 AGGTGATGTGTCCAAATCTAGATCCCAACTTCTCATTTCAAGGAGGAGGAGGAGGAGG 786
QY 926 CGCGGTGCCACTGGCTTTTATATGGAATGATGATGATGATGATGATGATGATGATGATG 985
Db 787 GGAACATGTGTGCTGGCATTATACAGCGGCTCTTTGCTGATGAGGAGGAGGAGGAGGAG 846
QY 986 ACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAACCCCTTCCCTTCCCAATATATATAT 1045
Db 847 ATTTGCTCACAGAGGAAATGATCAACCCCTACAGAAACCCCTTCCCTTCCCAATATATAT 906
QY 1046 CCATGGCCATTGTCACCATTTGGCTATGCTGACAAATGTGGCCTACTTTACGACCATTA 1105
Db 907 CCCTGCCCATCGTGACGCTGGTGTACGTGCTGACCAACCTGGCCTACTTCCACCCTGT 966

QY	1064	TTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGC	1123
Db	55	TCGGCTATGTGCTGACAAATGTGGCCTACTTTACNACCATTAAATGCTGAGGAGCTGCTGC	114
QY	1124	TTTCAAAATGCAGTGGCAGTGACCTTTCTGAGCGGTACTGGGAAATTTCTCATTAGCAG	1183
Db	115	TTTCAAAATGCAGTGGCAGTGACCTTTCTGAGCGGTACTGGGAAATTTCTCATTAGCAG	174
QY	1184	TTCCGATCTTTGTTGCCCTCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCT	1243
Db	175	TTCCGATCTTTGTTGCCCTCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCT	234
QY	1244	CCAGGTTATTTATGTTGGCTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTC	1303
Db	235	CCAGGTTATTTATGTTGGCTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTC	294
QY	1304	ATGTCGCGAAGCACACTCCTCTACCACTGTTATTGTTTTCACCCCTTTGACAAATGATAA	1363
Db	295	ATGTCGCGAAGCACACTCCTCTACCACTGTTATTGTTTTCACCCCTTTGACAAATGATAA	354
QY	1364	TGCTCTTCTCTGAGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTT	1423
Db	355	TGCTCTTCTCTGAGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTT	414
QY	1424	TTATTTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCAGATATGCATC	1483
Db	415	TTATTTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCAGATATGCATC	474
QY	1484	GTCTTTCAAGGTGGCACTGTTTCATCCCACTTTGTTTCCCTCAGTTTCCCTCAGTTTGG	1543
Db	475	GTCTTTCAAGGTGGCACTGTTTCATCCCACTTTGTTTCCCTCAGTTTCCCTCAGTTTGG	533
QY	1544	TTGCGCTTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGA	1603
Db	534	TTGCGCTTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGA	593
QY	1604	CTGGAGTCCCTGCGTATTATCTCTTTATATATGGGACAAGAAACCCAGGTGGTTTAGAA	1663
Db	594	CTGGAGTCCCTGCGTATTATCTCTTTATATATGGGACAAGAAACCCAGGTGGTTTAGAA	653
QY	1664	TAATGTCAGAGAAATAA-CCAGAACATTACAAATAATACTGGAAGTTGTACCAGAAGAA	1722
Db	654	TAATGTCAGAGAAATAAACCCGAACATTACAAATAATACTGGAAGTTGT-CCAGAAGAA	712
QY	1723	GATAAGTTATGAACATAATGGACTTGAGATCTTTGGCAATCTGCCAAGGGGAGACACAAA	1782
Db	713	GATAA-TTATGAACATAATGGACTTGAGA-CTTGGCAATCTG-CCAAGGGGAGACACAAA	769
QY	1783	TA 1784	
Db	770	TA 771	
RESULT 6			
AAZ17094			
ID	AAZ17094 standard; cDNA; 761 BP.		
XX			
AC	AAZ17094;		
XX			
DT	12-OCT-1999 (first entry)		
XX	Human gene expression product cDNA sequence SEQ ID NO:4565.		
DE			
XX	Human; gene; gene expression product; diagnosis; therapy; probe;		
KW	detection; mapping; tissue typing; profiling; forensic; cancer;		
KW	genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.		
XX	Homo sapiens.		
OS			
XX	WO9938972-A2.		
PN			
XX	05-AUG-1999.		
PD			
XX			

PF

XX

PR

PR

PR

PR

PR

PR

XX

PA

PA

XX

PI

PI

PI

PI

PI

XX

DR

XX

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

28-JAN-1999;

99WO-US01619.

03-APR-1998;

98US-0080666.

28-JAN-1998;

98US-0072910.

24-FEB-1998;

98US-0075954.

31-MAR-1998;

98US-0080114.

03-APR-1998;

98US-0080515.

(CHIR) CHIRON CORP.

(HYSE-) HYSEQ INC.

Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;

Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;

Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;

Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

Stache-Crain B, Sudduth-Klinger J, Williams LT;

WPI; 1999-494092/41.

Novel human genes and their expression products which are

differentially expressed in different cell types

Claim 1; Page 2164-2165; 2479pp; English.

The present invention describes a library of human polynucleotides

comprising the sequences given in AAZ12532 to AAZ17779. Also described is

a method of detecting differentially expressed genes correlated with the

cancerous state of a mammalian cell, comprising detecting at least one

differentially expressed gene product in a test sample from a cell

suspected of being cancerous, where the gene product is encoded by one

of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The

polynucleotides can be used as a source of primers and probes, which can

be used for a variety of purposes, e.g. detection of expression levels,

mapping, tissue typing or profiling, forensics, genetic analysis and

detection of polymorphisms. Polypeptides encoded by the polynucleotides

can be used for raising antibodies for experimental, diagnostic and

therapeutic purposes. The polynucleotides may also be used to construct

arrays for diagnostics (which may be used to determine function of an

encoded protein); and to detect differences in expression levels between

two cells (e.g. to identify abnormal or diseased tissue in a human, to

identify a genetic predisposition or susceptibility to a disease such as

cancer). The polynucleotides of the invention are especially used in the

diagnosis, prognosis and management of colorectal cancer, breast cancer,

and lung cancer. The polynucleotides can also be used to screen for

peptide analogues and antagonists.

Sequence 761 BP; 160 A; 162 C; 150 G; 250 T; 39 other;

Query Match

Best Local Similarity

Matches

641;

Conservative

0;

Mismatches

72;

Indels

14;

Gaps

8;

Length

761;

QY

805

TTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCAG

864

Db

18

TTTGGATCTNTGTGAGGATCCCATCGATTCTTAAAGACCGCTTTTCAGGAAGAGATTCAAGTNTT

75

QY

865

CTAATTAAGGTCAAACCGAGAACTTTAAAGACCGCTTTTCAGGAAGAGATTCAAGTNTT

924

Db

76

CTAATTAAGGTCAAACCGATTAACCTTTAAAGACCGCTTTTCAGGAAGAGATTCAAGTNTT

135

QY

925

ACGCGGTTGCCACTGGCTTTTATTATGGAATGATGATATGCTGGCTGGTTTACCTC

984

Db

136

ACGCGGTTGCCACTGGCTTTTATTATGGAATGATGATATGCTGGCTGGTTTACCTC

195

QY

985

AACCTTTGTTACTGAAGAAGTAGAAACCCCTGAAACCAACCATTTCCCTTGCATATGTATA

1044

Db

196

AACCTTTGTTACTGAAGAAGTAGAAACCCCTGAAACCAACCATTTCCCTTGCATATGTATA

255

QY

1045

TCCATGGCATTTGTCACCATTTGGCTATGCTGACAAATGTGGCCTACTTACGACCATT

1104

Db

256

TCCATGGCATTTGTCACCATTTGGCTATGCTGACAAATGTGGCCTACTTACGACCATT

315

QY

1105

AATGCTGAGGAGCTGCTGCTTTCAAATGCA-GTGGCACTGACCTTTTCTGAGGGGCTACT

1163

QY	1072	GTGCTGACAAATGTGGCCTACTTTACGACCATTAACTGCTGAGGAGCTGCTGCTTCAAAT	1131
Db	906	GTGCTGACAAATGTGGCCTACTTTACGACCATTAACTGCTGAGGAGCTGCTGCTTCAAAT	965
QY	1132	GCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATC	1191
Db	966	GCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATC	1025
QY	1192	TTTGTGGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTTGGTCTCCAGGTTA	1251
Db	1026	TTTGTGGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTGTGTTTGGTCTCCAGGTTA	1085
QY	1252	TTCTATGTTGGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCCGC	1311
Db	1086	TTCTATGTTGGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCCGC	1145
QY	1312	AAGCACACTCCTCTACAGCTGTTATTGTTTGCACCCTTTTGACAATGATAATGCTCTTC	1371
Db	1146	AAGCACACTCCTCTACAGCTGTTATTGTTTGCACCCTTTTGACAATGATAATGCTCTTC	1205
QY	1372	TCGGGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTTCGCCAGGTGGCTTTTATTTGGG	1431
Db	1206	TCGGGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTTCGCCAGGTGGCTTTTATTTGGG	1265
QY	1432	CTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTTC	1491
Db	1266	CTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTTC	1325
QY	1492	AAGGTGCCACTGTTCATCCAGCTTTTGTTCCTTCACATGCCTCTTCATGGTGGCCTT	1551
Db	1326	AAGGTGCCACTGTTCATCCAGCTTTTGTTCCTTCACATGCCTCTTCATGGTGGCCTT	1385
QY	1552	TCCCTCTATTCCGACCCCATTTAGTACAGGGATTGGCTTCCTCATCTCATCTGAGTGGAGTC	1611
Db	1386	TCCCTCTATTCCGACCCCATTTAGTACAGGGATTGGCTTCCTCATCTCATCTGAGTGGAGTC	1445
QY	1612	CCTGCGTATTATCTCTTTTATTATATGGGACAAGAACCCAGGTGGTTTAGAATAATGTCA	1671
Db	1446	CCTGCGTATTATCTCTTTTATTATATGGGACAAGAACCCAGGTGGTTTAGAATAATGTCA	1505
QY	1672	GAGAAATAACAGAACATTTACAAATAACTGGAAGTTGTACCAGAGAGATAAGTTA	1731
Db	1506	GAGAAATAACAGAACATTTACAAATAACTGGAAGTTGTACCAGAGAGATAAGTTA	1565
QY	1732	TGAACATAATGGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAAATAGGATTT	1791
Db	1566	TGAACATAATGGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAAATAGGATTT	1625
QY	1792	TTACTTCATTTTCTGAAAGTCTAGAGAATTACAACTTTGGTGATAAACAAGGAGTCA	1851
Db	1626	TTACTTCATTTTCTGAAAGTCTAGAGAATTACAACTTTGGTGATAAACAAGGAGTCA	1685
QY	1852	TTATTTTATTTCATATATTTTAGCATATTTTGAAGTATTTCTAGTGCACAAATTTCTTGAGTC	1911
Db	1686	TTATTTTATTTCATATATTTTAGCATATTTTGAAGTATTTCTAGTGCACAAATTTCTTGAGTC	1745
QY	1912	TCATGTAGTTATAGAAAGTGAATATGCAATATTTCTAGTGCACAAATTTCTTGAGTC	1971
Db	1746	TCATGTAGTTATAGAAAGTGAATATGCAATATTTCTAGTGCACAAATTTCTTGAGTC	1805
QY	1972	TCTGATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCATCT	2031
Db	1806	TCTGATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCATCT	1865
QY	2032	CTACAACATATGTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTAT	2091
Db	1866	CTACAACATATGTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTAT	1925
QY	2092	ATATGGGTTTGTAAAGATGGTTTACACACTA	2124
Db	1926	ATATGGGTTTGTAAAGATGGTTTACACACTA	1958

RESULT 5

AAZ16528
ID AAZ16528 standard; cDNA; 772 BP.

XX
AC AAZ16528;
XX
DT 12-OCT-1999 (first entry)

XX Human gene expression product cDNA sequence SEQ ID NO:3998.

Human; gene; gene expression product; diagnosis; therapy; probe;
detection; mapping; tissue typing; profiling; forensic; cancer;
genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

OS Homo sapiens.

XX WO9938972-A2.

XX 05-AUG-1999.

XX 28-JAN-1999; 99WO-US01619.

PR 03-APR-1998; 98US-0080666.

PR 28-JAN-1998; 98US-0072910.

PR 24-FEB-1998; 98US-0075954.

PR 31-MAR-1998; 98US-0080114.

PR 03-APR-1998; 98US-0080515.

XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.

PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX WPI; 1999-494092/41.

PT Novel human genes and their expression products which are
PT differentially expressed in different cell types

PS Claim 1; Page 1897; 2479pp; English.

XX The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.

SQ Sequence 772 BP; 175 A; 183 C; 158 G; 248 T; 8 other;

Query Match

Best Local Similarity 28.8%; Score 643.8; DB 20; Length 772;

Matches 711; Conservative 0; Mismatches 5; Indels 6; Gaps 6;

giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

SQ Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;

Query Match	81.4%	Score 1822.2;	DB 22;	Length 1958;
Best Local Similarity	95.4%	Pred. No. 0;		
Matches 1920; Conservative	0;	Mismatches	38; Indels	55; Gaps
				2;

QY	112	CAATTGTGGCATAGATATTTTATCATATCTCGATTTTTTGGAATCTTTGTTTCTCTCATCA	171
Db			
	1	CAATTCGGGTCGACGATTCGTTCCTCTGTTTCCCTCTGTTTATATTTTCCCCCGTGTGCCCTA	60
QY	172	CTGGATTCAGGAAAGCCTGTTGTGTCCACCATCTCCAAGAGAGGTTACCTGCAGGGAAT	231
Db			
	61	CTATGGTCA-GAAAGCCTGTTGTGTCCACCATCTCCAAGAGAGGTTACCTGCAGGGAAT	119
QY	232	GTTAACGGGAGGCTGCCTTCCCCTGGGCAACAAGGAGCACCTGGGCAGGACGCTTTTCA	291
Db			
	120	GTTAACGGGAGGCTGCCTTCCCCTGGGCAACAAGGAGCACCTGGGCAG-----	167
QY	292	GGAAGAGACGCCCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAAGTCAGCTGAAG	351
Db		-----GAGAAAGTCAGCTGAAG	168
QY	352	AGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGA	411
Db			
	186	AGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGA	245
QY	412	ATCTTCATCTCTCCTAAGGGCGTGTCTCCAGAACACGGCAGCGTGGSCATGTCTCTGACC	471
Db			
	246	ATCTTCATCTCTCCTAAGGGCGTGTCTCCAGAACACGGCAGCGTGGSCATGTCTCTGACC	305
QY	472	ATCTGGACGGTGTGTGGGTCTCTGTCACTAHTTGGAGCTTTGTCTTATGCTGAATTGGGA	531
Db			
	306	ATCTGGACGGTGTGTGGGTCTCTGTCACTAHTTGGAGCTTTGTCTTATGCTGAATTGGGA	365
QY	532	ACAACTATAAAGAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTTA	591
Db			
	366	ACAACTATAAAGAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTTA	425
QY	592	CCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTG	651
Db			
	426	CCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTG	485
QY	652	ATATCCCTGGCAATTTGGACGCTACAHTCTGGAACCAATTTTATTTCAATGTGAATCCCT	711

QY	903	TTTACGAGAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATATATGGAATGTATGC	962
Db	847	TTTACGAGAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATATGGAATGTATGC	906
QY	963	ATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCTGAAAAAC	1022
Db	907	ATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCTGAAAAAC	966
QY	1023	CATTCCCTTGCATATATATCCATGGCCATTGTCACCAATTGGCTATGTGCTGACAAA	1082
Db	967	CATTCCCTTGCATATATATCCATGGCCATTGTCACCAATTGGCTATGTGCTGACAAA	1026
QY	1083	TGTGGCCTACTTTACGACCAATTAAATGCTGAGGAGCTGCTGCTTTCAAATGCAAGTGGCAGT	1142
Db	1027	TGTGGCCTACTTTACGACCAATTAAATGCTGAGGAGCTGCTGCTTTCAAATGCAAGTGGCAGT	1086
QY	1143	GACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTTCCGATCTTTTGTGCCCCT	1202
Db	1087	GACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTTCCGATCTTTTGTGCCCCT	1146
QY	1203	CTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTTCTATGTTGC	1262
Db	1147	CTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTTCTATGTTGC	1206
QY	1263	GTCTCGAGAGGGTCCACTTCCAGAAATCCTCTCCATGATTCTCCGCAAGCACACTCC	1322
Db	1207	GTCTCGAGAGGGTCCACTTCCAGAAATCCTCTCCATGATTCTCCGCAAGCACACTCC	1266
QY	1323	TCTACCAAGCTGTTTATTTGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCT	1382
Db	1267	TCTACCAAGCTGTTTATTTGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGACCT	1326
QY	1383	CGACAGCTTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTTTATTTATGGGCTGGCAGTTGC	1442
Db	1327	CGACAGCTTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTTTATTTATGGGCTGGCAGTTGC	1386
QY	1443	TGGGCTGATTTATCTTCGATACAAATGCCCCAGATATGCATGCTCCTTTCAAGGTGCCACT	1502
Db	1387	TGGGCTGATTTATCTTCGATACAAATGCCCCAGATATGCATGCTCCTTTCAAGGTGCCACT	1446
QY	1503	GTTTCATCCAGCTTTGTTTTCCTTCACATGCTTCTCATGTTTGGCCCTTCCCTCTATTC	1562
Db	1447	GTTTCATCCAGCTTTGTTTTCCTTCACATGCTTCTCATGTTTGGCCCTTCCCTCTATTC	1506
QY	1563	GGACCCATTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTA	1622
Db	1507	GGACCCATTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTA	1566
QY	1623	TCTCTTTATATATGGGACAGAAACCCAGGTGGTTTGAATAATGTCAGAGAAATAAC	1682
Db	1567	TCTCTTTATATATGGGACAGAAACCCAGGTGGTTTGAATAATGTCAGAGAAATAAC	1626
QY	1683	CAGAACATTACAAATAATACTGGAAGTTGTACCAGAGAGACACAAATAGGGATTTTACTTCATT	1742
Db	1627	CAGAACATTACAAATAATACTGGAAGTTGTACCAGAGAGACACAAATAGGGATTTTACTTCATT	1686
QY	1743	ACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCATT	1802
Db	1687	ACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCATT	1746
QY	1803	TCTGAAAGTCTAGAGAAATTACAACCTTTGGTGATAAACAAGAGAGTCAAGTTATTTTATT	1862
Db	1747	TCTGAAAGTCTAGAGAAATTACAACCTTTGGTGATAAACAAGAGAGTCAAGTTATTTTATT	1806
QY	1863	CATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATAACTCTATGAGTT	1922
Db	1807	CATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATAACTCTATGAGTT	1866
QY	1923	ATAGAAAGTGAATATGAGTTATTTCTATGAGTCGCACAATTTCTGAGTCTCTGATACCTA	1982
Db	1867	ATAGAAAGTGAATATGAGTTATTTCTATGAGTCGCACAATTTCTGAGTCTCTGATACCTA	1926
QY	1983	CCTATTGGGTTAGGAGAAAAGACTAGACAATTTACTATGTGTCATTTCTCTACACATAT	2042
Db	1927	CCTATTGGGTTAGGAGAAAAGACTAGACAATTTACTATGTGTCATTTCTCTACACATAT	1986
QY	2043	GTTAGCACGGCAAGAACCTTCAAAATTGAAGACTGAGATTTTCTGTATATATGGGTTT	2102
Db	1987	GTTAGCACGGCAAGAACCTTCAAAATTGAAGACTGAGATTTTCTGTATATATGGGTTT	2046
QY	2103	GTAAAGATGGTTTTTACACACTACAGATGTCTATACTGTGAAAAGTGTTTTCAATTCGAA	2162
Db	2047	GTAAAGATGGTTTTTACACACTACAGATGTCTATACTGTGAAAAGTGTTTTCAATTCGAA	2106
QY	2163	AAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAAGTA	2204
Db	2107	AAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAAGTA	2148
RESULT 3			
ABA09201			
ID	ABA09201	standard; cDNA; 1958 BP.	
XX	ABA09201;		
AC	ABA09201;		
XX	11-JAN-2002	(first entry)	
DT	11-JAN-2002	(first entry)	
XX	Human cystine/Glu transporter homologue cDNA, SEQ ID NO:977.		
DE	Human; cytokine; cell proliferation; cell differentiation; growth factor;		
XX	haematopoiesis regulation; tissue growth; immunomodulator; activin;		
KW	inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;		
KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;		
KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;		
KW	chronic inflammatory condition; proliferative retinopathy;		
KW	atherosclerosis; coronary heart disease; arterial ischaemia;		
KW	bone disorder; osteoporosis; vascular growth disorder;		
KW	tissue regeneration; wound healing; infection; immune disorder;		
KW	cell culture; drug screening; gene therapy; antiinflammatory;		
KW	antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;		
KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;		
KW	antifungal; vulnery; antiulcer; ss.		
XX	Homo sapiens.		
OS	WO200157188-A2.		
XX	09-AUG-2001.		
PN	05-FEB-2001; 2001WO-US03800.		
XX	03-FEB-2000; 2000US-0496914.		
XX	27-APR-2000; 2000US-0560875.		
PF	(HYSE-) HYSEQ INC.		
XX	Tang YT, Liu C, Drmanac RT;		
PD	WPI; 2001-457740/49.		
XX	P-PSDB; ABB11957.		
XX	Human proteins and DNA encoding sequences useful for preventing,		
XX	treating or ameliorating a medical condition in a mammalian subject		
XX	e.g. arthritis and cancer -		
XX	Claim 1; Page 835; 1963pp; English.		
XX	Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and		
CC	sequences ABA08225-ABA09574 represent nucleic acids encoding them. The		
CC	invention also relates to vectors and recombinant host cells comprising a		
CC	nucleotide of the invention, methods of producing the novel polypeptides,		
CC	antibodies against the polypeptides, methods of detecting the nucleotides		
CC	or polypeptides in a sample, and methods of identifying compounds which		
CC	bind to polypeptides of the invention. Although novel, many of the		
CC	polypeptides of the invention have homology to known proteins, thereby		

Db	1801	TTTCTGAAAGTCTAGAGAAATTACAACCTTTGGTGATAAACAAAGGAGTCAGTTATTTTA	1860
QY	1861	TTCATATATTTTAGCATATTCGAACTAAATTTCTAAGAAATTTAGTTATAACTCTATGTAG	1920
Db	1861	TTCATATATTTTAGCATATTCGAACTAAATTTCTAAGAAATTTAGTTATAACTCTATGTAG	1920
QY	1921	TTATAGAAAGTGAATATGCGAGTTATCTATGAGTCGCACAAATCTTGAGTCTCTGATACC	1980
Db	1921	TTATAGAAAGTGAATATGCGAGTTATCTATGAGTCGCACAAATCTTGAGTCTCTGATACC	1980
QY	1981	TACCTATTGGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGGTCATTTCTCTACAACAT	2040
Db	1981	TACCTATTGGGGTTAGGAGAAAAGACTAGACAAATTACTATGTGGTCATTTCTCTACAACAT	2040
QY	2041	ATGTTAGCACGGCAAAGAACCCTTCAAATTGAACACTGAGATTTTCTGTATATATGGGTT	2100
Db	2041	ATGTTAGCACGGCAAAGAACCCTTCAAATTGAACACTGAGATTTTCTGTATATATGGGTT	2100
QY	2101	TTGTAAGATGGTTTACACACTACAGATGCTATACTGTGAAAAGTGTTCAAATTCGTG	2160
Db	2101	TTGTAAGATGGTTTACACACTACAGATGCTATACTGTGAAAAGTGTTCAAATTCGTG	2160
QY	2161	AAAAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAGGTAGAGCTGTTCTTAAAT	2220
Db	2161	AAAAAAGCATACATCATGATTATGGCAAAGAGGAGAGAAGGTAGAGCTGTTCTTAAAT	2220
QY	2221	TATTAATAAAAAAAAAAAAAA	2239
Db	2221	TATTAATAAAAAAAAAAAAAA	2239

RESULT 2

AAD23461

ID	AAD23461	standard	CDNA	5981	BP
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

XX
XX

AC AAD23461;

XX

DT 26-FEB-2002 (first entry)

XX
XX
(Transmitted)

DE Human lung tu

XX

Human; lung

antisen-therapy; vaccine; vaccine; vaccine

1. **Introduction**
 2. **Background**
 3. **Methods**
 4. **Results**
 5. **Conclusion**
 6. **References**

US
yy
Homo sapiens.

XX
PN
W030017230E 23

FN
XX
W02001/2295-A2.

04-OCT-2001

XX
XX
1002-130 45

28-MAR-2001: 2001WO-
PFXX
OMT002 'T002 www.3

PR 29-MAR-2000: 2000US-

PR 05-JUN-2000; 2000US-

PR 18-AUG-2000; 2000US-0640878.

PR 22-SEP-2000; 2000US-234517P.

PR 01-NOV-2000; 2000US-0704512.

PR 14-DEC-2000; 2000US-0738973.

XX	DR	WPI; 2001-639201/73.	
XX	PT	New human lung-specific polynucleotides and polypeptides for the diagnosis and treatment of disease e.g. lung cancer -	
XX	PS	Claim 1; Page 332; 378pp; English.	
XX	CC	The invention relates to isolated lung tumour-specific proteins and their corresponding cDNA molecules. Lung tumour-specific proteins and their antigen-presenting cells are useful for stimulating and/or expanding T cells specific for a tumour protein, and for inhibiting the development of cancer. The invention also relates to a composition useful for stimulating an immune response, and for treating cancer. The lung tumour specific oligonucleotide is useful in gene therapy and for diagnosis, detection and treatment of lung cancer. The present sequence is a cDNA encoding human lung tumour-specific protein.	
XX	SQ	Sequence 2239 BP; 619 A; 444 C; 493 G; 583 T; 0 other;	
		Query Match 100.0%; Score 2239; DB 23; Length 2239;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	Db	1 GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAC 60	
QY	Db	1 GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAC 60	
QY	Db	61 AGAATTAAGGAAAAAGAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCCAATTGTGG 120	
QY	Db	61 AGAATTAAGGAAAAAGAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCCAATTGTGG 120	
QY	Db	121 CATAGATTTTATCATATATCTGGATTTTGTGGATTTCTTTGTCTCATCATCTGGAATCA 180	
QY	Db	121 CATAGATTTTATCATATATCTGGATTTTGTGGATTTCTTTGTCTCATCATCTGGAATCA 180	
QY	Db	181 GGAAGCCTGTTGTGTCCACCATCTCCAAAGAGAGTTACCTGCAGGAAATGTTAACGGG 240	
QY	Db	181 GGAAGCCTGTTGTGTCCACCATCTCCAAAGAGAGTTACCTGCAGGAAATGTTAACGGG 240	
QY	Db	241 AGGCTGCCTTCCCTGGCAACAAGAGAGCCACCTGGGAGGAGGCTTTTCAGGAAGAGAC 300	
QY	Db	241 AGGCTGCCTTCCCTGGCAACAAGAGAGCCACCTGGGAGGAGGCTTTTCAGGAAGAGAC 300	
QY	Db	301 GCCTTTTCAGGAAGAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAATGC 360	
QY	Db	301 GCCTTTTCAGGAAGAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAATGC 360	
QY	Db	361 ACTTTACTGAGGGAGTCTCCATATATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATC 420	
QY	Db	361 ACTTTACTGAGGGAGTCTCCATATATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATC 420	
QY	Db	421 TCTCCTAAGGCGTCTCCAGAACACGGGAGCGTGGGCATGTCTCTGACCATCTGGAGC 480	
QY	Db	421 TCTCCTAAGGCGTCTCCAGAACACGGGAGCGTGGGCATGTCTCTGACCATCTGGAGC 480	
QY	Db	481 GTGTGTGGGGTCTGTCACTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATA 540	
QY	Db	481 GTGTGTGGGGTCTGTCACTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATA 540	
QY	Db	541 AAGAAATCTGGAGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAAGCTTTT 600	
QY	Db	541 AAGAAATCTGGAGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTACCAAGCTTTT 600	
QY	Db	601 GTACGAGTCTGGGTGGAACCTCCTCAATATACGCCCTGCAGCTACTGTGTGATATCCCTG 660	
QY	Db	601 GTACGAGTCTGGGTGGAACCTCCTCAATATACGCCCTGCAGCTACTGTGTGATATCCCTG 660	
QY	Db	661 GCAATTTGACGCTACATTTCTGGAACCAATTTTATCAATGTGAATTTCCCTGAACTTGGC 720	
QY	Db	661 GCAATTTGACGCTACATTTCTGGAACCAATTTTATCAATGTGAATTTCCCTGAACTTGGC 720	
QY	Db	721 ATCAAGCTCATACAGCTGTGGGCATAACTGTAGTGGTCTTAAATAGCATGAGTGC 780	

Db	721	ATCAAGCTCATACAGCTGTGGGCATAAAGTGTAGTGGTCTTAAATAGCATGAGTGC 780	
QY	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGCAGCTCACAGCAATTTCTGATA 840	
Db	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGCAGCTCACAGCAATTTCTGATA 840	
QY	841	ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACAAACGAGAACTTTAAAGACGCC 900	
Db	841	ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACAAACGAGAACTTTAAAGACGCC 900	
QY	901	TTTTCAGGAAGAGATTTCAAGTATTACCGGTTTCCACTGGCTTTTATTATGGAATGTAT 960	
Db	901	TTTTCAGGAAGAGATTTCAAGTATTACCGGTTTCCACTGGCTTTTATTATGGAATGTAT 960	
QY	961	GCATATGCTGGCTGGTTTTACCTCAACTTTGTACTGAAGAAAGTAGAAAAACCCCTGAAAAA 1020	
Db	961	GCATATGCTGGCTGGTTTTACCTCAACTTTGTACTGAAGAAAGTAGAAAAACCCCTGAAAAA 1020	
QY	1021	ACCATTCCTCTTGCATATATATCCATGGCCATTTGTCACCATTTGGCTATGTGCTGACA 1080	
Db	1021	ACCATTCCTCTTGCATATATATCCATGGCCATTTGTCACCATTTGGCTATGTGCTGACA 1080	
QY	1081	AATGTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCA 1140	
Db	1081	AATGTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGGCA 1140	
QY	1141	GTGACCTTTTCTGAGCGGCTACTTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1200	
Db	1141	GTGACCTTTTCTGAGCGGCTACTTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1200	
QY	1201	CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTT 1260	
Db	1201	CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTT 1260	
QY	1261	GGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCATGATTCATGTCGGAAGCACACT 1320	
Db	1261	GGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCATGATTCATGTCGGAAGCACACT 1320	
QY	1321	CCTCTACCAAGCTGTTATTTGTTTGGACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGAC 1380	
Db	1321	CCTCTACCAAGCTGTTATTTGTTTGGACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGAC 1380	
QY	1381	CTCGACAGCTTTTGAATTTTCCCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTT 1440	
Db	1381	CTCGACAGCTTTTGAATTTTCCCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTT 1440	
QY	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGGTGCCA 1500	
Db	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGGTGCCA 1500	
QY	1501	CTGTTTCATCCCAGCTTTGTTTCCCTTCACATGCCCTTCTCATGGTGGCTTTTCCCTCTAT 1560	
Db	1501	CTGTTTCATCCCAGCTTTGTTTCCCTTCACATGCCCTTCTCATGGTGGCTTTTCCCTCTAT 1560	
QY	1561	TCGGACCCATTTAGTACAGGGATTGGCTTCCCTTCACATCCTCTGACTGGAGTCCCTGCCGTAT 1620	
Db	1561	TCGGACCCATTTAGTACAGGGATTGGCTTCCCTTCACATCCTCTGACTGGAGTCCCTGCCGTAT 1620	
QY	1621	TATCTCTTTTATATATGGGACAAAGAAACCCAGGTGTTTAGAATAATGTCAGAGAAAAATA 1680	
Db	1621	TATCTCTTTTATATATGGGACAAAGAAACCCAGGTGTTTAGAATAATGTCAGAGAAAAATA 1680	
QY	1681	ACCAGAACATTTACAAATAATGCTGGAAGTTGTACACGAAAGAGATAAGTTATGAACATAAT 1740	
Db	1681	ACCAGAACATTTACAAATAATGCTGGAAGTTGTACACGAAAGAGATAAGTTATGAACATAAT 1740	
QY	1741	GGACTTGAGATCTTGGCAATCTGCCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCTAT 1800	
Db	1741	GGACTTGAGATCTTGGCAATCTGCCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCTAT 1800	
QY	1801	TTTCTGAAAGTCTAGAGAAATTTACAACCTTTGGTGATAAAACAAAGGAGTCAAGTTATTTTA 1860	

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 03:52:53 ; Search time 521 Seconds
(without alignments)
9677.974 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	100.0	2239	23	Human lung tumour-
2	2080.8	92.9	5981	23	Human lung tumour-
3	1822.2	81.4	1958	22	Human cysteine/Glu
4	1822.2	81.4	1958	22	Human EST-derived
5	643.8	28.8	772	20	Human gene express
6	520.6	23.3	761	20	Human gene express
7	370.8	16.6	4539	21	Human L-type amino
8	366.4	16.4	1533	21	Human monp-2 cDNA.
9	366.4	16.4	1656	21	Human monp-2 cDNA

10	366.4	16.4	2072	20	AAZ23240	Human amino acid p
11	366.4	16.4	2072	20	AAZ23240	Human amino acid p
12	362.2	16.2	1524	21	AAA08393	Human L-type amino
13	353.6	15.8	4585	22	ABA09358	Human 4F2 light ch
14	344.8	15.4	4117	22	AAF77276	Rat cDNA encoding
15	343.8	15.4	1632	21	AAZ55789	cDNA encoding rat
16	343.8	15.4	2155	21	AAZ00605	Human membrane tra
17	343.8	15.4	4211	22	AAF77278	Human cDNA encodin
18	343.8	15.4	4261	21	AAZ77345	Human ORFX ORF2900
19	342.2	15.3	3455	21	AAA08373	Rat L-type amino a
20	331.6	14.8	1716	22	AAH74767	Nucleotide sequenc
21	329.2	14.7	2476	23	ABL28785	Drosophila melanog
22	328.4	14.7	2090	23	ABL10465	Drosophila melanog
23	325	14.5	1997	23	ABL02951	Drosophila melanog
24	315.6	14.1	1918	22	AAH74769	Nucleotide sequenc
25	311.4	13.9	1554	23	ABL12149	Drosophila melanog
26	308.2	13.8	800	20	AAZ16609	Human gene express
27	303	13.5	596	24	ABK27651	Human colon cancer
28	301	13.4	1461	22	AAF94439	Human hydrophobic
29	301	13.4	1758	22	AAF94449	Human hydrophobic
30	300	13.4	300	20	AAZ14980	Human gene express
31	300	13.4	300	20	AAZ14962	Human gene express
32	288	12.9	288	20	AAZ14304	Human gene express
33	287.8	12.9	300	21	AAA00880	Human colon cancer
34	286.2	12.8	3735	22	ABA09032	Human amino acid t
35	281	12.6	337	23	AAZ23462	Human lung tumour-
36	260.8	11.6	1551	23	ABL15177	Drosophila melanog
37	238.4	10.6	3984	24	ABK84083	Human cDNA differe
38	238.4	10.6	3984	24	ABL62111	Colon adenocarcino
39	238.4	10.6	3984	24	ABL62881	Breast cancer rela
40	230	10.3	575	24	ABL37718	Human colon tumour
41	229	10.2	1564	21	AAZ26346	Human secreted pro
42	226.8	10.1	2256	21	AAZ99060	Human pancreatic c
43	224	10.0	1161	23	AAZ76312	DNA encoding novel
44	215.2	9.6	1397	24	ABL90625	Human polynucleoti
45	211.4	9.4	1011	19	AAV68500	Human tumour-assoc

ALIGNMENTS

RESULT 1
AAD23460
ID AAD23460 standard; cDNA; 2239 BP.

XX AAD23460;

AC AAD23460;

XX AAD23460;

DT 26-FEB-2002 (first entry)

XX Human lung tumour-specific 19A4 cDNA.

DE Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;

XX antisense-therapy; vaccine; immune response; lung cancer; 19A4; ss.

OS Homo sapiens.

XX WO200172295-A2.

PN 04-OCT-2001.

XX 28-MAR-2001; 2001WO-US09991.

XX 29-MAR-2000; 2000US-0538037.

XX 05-JUN-2000; 2000US-0588937.

XX 18-AUG-2000; 2000US-0640878.

XX 22-SEP-2000; 2000US-234517P.

XX 01-NOV-2000; 2000US-0704512.

XX 14-DEC-2000; 2000US-0738973.

XX (CORI-) CORIXA CORP.

XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;

PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;

QY 242 GGCTGCCTTCCCTGGGCAACAAGGAGGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGACG 301
Db 198 GGCTCCCTCCCTGGGAGCAAGAGCCACCTGGGCAAT-----235
QY 302 CCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAGTCA 361
Db 236 -----GAGAAGGTGGTCTTGAAAAAGAGATCA 263
QY 362 CTTTACTGAGGGAGTCTCCATATATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT 421
Db 264 CTTTGTGAGGGGGTCTCCATCATCATCGGACCGCTCACTCGGATCGGCGCATCTTCATCT 323
QY 422 CTCCTAAGGGCGTCTCCAGAACACGGGCGGCGTGGGCGATGTCTCTTGACCATCTGGACGG 481
Db 324 CCCCCAAGGGCATACTCCAGAACACGGGCGGCGTGGGCGATGTCACTGGTGTCTGGTCTG 383
QY 482 TGTGTGGGGTCTGTCACTATTATTGGAGCTTTGTCTTATGTGAATTGGGAACAATAA 541
Db 384 CTTGTGGAGTACTGTCACTTTTGGAGCCCTGTCTTATGTGAATTGGGTACGAGCATAA 443
QY 542 AGAAATCTGGAGTCAATTACACATATATTTTGGAAAGTCTTGGTCCATTACAGCTTTTG 601
Db 444 AGAAATCTGGTGTCAATTACACATATATTCTGGAGGCTTTGGTCCCTTGCTAGCTTTG 503
QY 602 TACGAGTCTGGGTGGAACCTCC 622
Db 504 TTCGAGTCTGGGTGGAACCTGC 524

RESULT 15
BQ684696
LOCUS BQ684696 924 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8033076 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091235
5', mRNA sequence.
ACCESSION BQ684696
VERSION BQ684696.1 GI:21797375
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 924)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/Drp
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2333 row: 0 column: 12
High quality sequence stop: 692.
Location/Qualifiers
1..924
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6091235"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 157 a 312 c 226 g 228 t 1 others
FEATURES
source

ORIGIN

Query Match 9.9%; Score 222.6; DB 14; Length 924;
Best Local Similarity 54.9%; Pred. No. 9.8e-33;
Matches 463; Conservative 0; Mismatches 374; Indels 6; Gaps 1;

QY 789 CGCCCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATAATATAGT 848
Db 2 CACCCGGGTCCAGGATGCCCTTTGCCCGCCCAAGCTCTGGCCCTGGCCCTGATCATCCT 61
QY 849 CCCTGGAGTTATGCAGCTAATTAAAGTCAAACGCAAGAACTT-----TAAAGACGCCTT 902
Db 62 GCTGGGCTTCGTCCAGATCGGAAGGTGATGTCCAATCTAGATCCCAACTTCTCAT 121
QY 903 TTCAGGAAGAGATTCAAGTATTACGGGTTGGCACTGGCTTTTATTTATTTATGGAATGATGC 962
Db 122 TGAAGGCACCAAACTGGATGTGGGAACAATTTGTCTGGCAATTATACAGCGGCTCTTTGC 181
QY 963 ATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAAGTAGAAAAACCTGAAAAAAC 1022
Db 182 CTATGGAGGATGGAATTAATTTGTCACAGAGAGAAATGATCAACCCCTACAGAAA 241
QY 1023 CATTCCTCTTGCATATATATCCATGGCCATTTGACCATTTGCTGCTGATGCTGACAAA 1082
Db 242 CTTGGCCCTGGCCATCATCTCCCTGGCCATCTGACGCTGCTGCTGCTGCTGCTGCTGCT 301
QY 1083 TGTGGCCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTCAATGCAATGCAATG 1142
Db 302 CTTGGCCTACTTCAACCACTGTCCACCGAGCAGATGCTGCTCGGAGGCGCTGGCCGT 361
QY 1143 GACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTTAGCAGTCCGATCTTTGTTGCCCT 1202
Db 362 GGACTTCGGGAATCATCACCTGGGCGTCATGCTCTGATCATCCCCCTCTTCGTTGGCCT 421
QY 1203 CTCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTGCTGCTCCAGGTTATTCATGTTGC 1262
Db 422 GTCTGCTTCCGCTCCGTCAATGGTTCCTGTTCACATCTCTCCAGGCTCTTCTTCGTGGG 481
QY 1263 GTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGCAAGCACTCC 1322
Db 482 GTCCCGGGAAGGCCACCTGCCCTCCATCTCTCCATGATCCACCCACAGCTCCTCACCCC 541
QY 1323 TCTACGAGCTGTATTGTTTGGCACCCTTTTGACAATGATAATGCTCTTCTCTGGAGACCT 1382
Db 542 CGTGCCGCTCCCTCGTGTTCACGTTGTGATGACGCTGCTCTACGCTTCTCCAAGGACAT 601
QY 1383 CGACAGTCTTTGAATTTCTCCTCAGTTTGGCCAGGTTGGCTTTTATTGGGCTGGCAGTTGC 1442
Db 602 CTTCTCCGTCATCAACTTCTCAGCTTCTTCAACTGGCTCTGCGTGGCCCTGGCCATCAT 661
QY 1443 TGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGATCGTCTTCAAGGTGCCACT 1502
Db 662 CGGCATGATCTGGCTGGCGCCACAGAAAGCCTGAGCTTGAGCGGCCCATCAAGGTGAACCT 721
QY 1503 GTTCATCCCGAGCTTTGTTTCTTCCATTCACATGCCCTTTCATGTTGGCTTCCCTCTATTC 1562
Db 722 GGCCCTGCCTGTGTTCTTCATCTCTGGGCTGGCTTCTCTGATCGCCCTCTCTCTTCTG 781
QY 1563 GGACCCATTTAGTACAGGATTTGGCTTCCGTCATCACTCTGACTGGAGTCCCTGCGTATTA 1622
Db 782 GACACCCGTGGAGTGTGGCATCGGCTTCAACCATCATCTCTCAGCGGCTGCCGTCACAACT 841
QY 1623 TCT 1625
Db 842 TCT 844

Search completed: April 16, 2003, 09:16:21
Job time : 3562 secs

[illegible]

QY	1405	AGTTTGGCCAGGTGGCTTTTATTATGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATAC	1464
Db	1320	TCC TTCATCAACTACCTCTGCTACGGCGTGACAATCCTGGGCCTGCTGCTGCTGCGCTGG	1379
QY	1465	AAATGCCCAGATATGCATCGTCCCTTTC AAGGTGCCACTGTTTCATCCCAGCTTGTGTTTCC	1524
Db	1380	AGCGGGCTGCAC TCCACAGGCCCATCAAGGTGAACCTTCTCATCCCCGTGGGTA CT TG	1439
QY	1525	TTACATAGCCTCTTCATGGTTGCCCTTTCCCTCTATT CGGACCATT TAGTACAGGGATT	1584
Db	1440	GTCTTCTGGGCCCTTCCTGCTGGTCTTCAGCTTCATCTCAGAGCCATG GTC TCTGGGGT C	1499
QY	1585	GGCTTCGT CATCACTCTGACTGGAGTCCC TGGGTATTATCTCTTTATATATGGGACAAG	1644
Db	1500	GGCGTCATCATCCTTACGGGGGTGCCCATTTTCTTCTGGGAGTGTTCTCTGGAGAAGC	1559
QY	1645	AAACCCAGGTGGTTTTAGAAATAATGTCAGAGAAAAAATAC	1682
Db	1560	AAACCAAAGTGTGTGCACAGACTCACAGAGTCCATGAC	1597

RESULT 11					
BJ075046	BJ075046	403 bp	mRNA	linear	EST 11-DEC-2001
LOCUS	BJ075046	NIBB Mochii normalized xenopus tailbud library	xenopus	tailbud library	xenopus
DEFINITION	laevis cDNA clone XL087e10 5', mRNA sequence.				

REFERENCE AUTHORS	1 (bases 1 to 403) Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara ,Y.
TITLE	Expressed genes in <i>X. laevis</i> embryo
JOURNAL	Unpublished (2001)
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.

```

FEATURES
source
Location/Qualifiers
1. .403
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL087e10"
/clone_lib="NIBB Mochii normalized xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
91 a 84 c 87 g 139 t 2 others
BASE COUNT
ORIGIN

```

	Query Match	11.2%;	Score 250;	DB 13;	Length 403;	
	Best Local Similarity	78.0%;	Pred. No. 7.7e-38;			
	Matches 312; Conservative	0;	Mismatches 87;	Indels	1;	Gaps
						1
QY	919 AGTATTACGGGTTGCCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTT	978				
Ddb	4 AGTGTTATGGGGTTACCCCTGGCATTCTATTCTGGAATGTATGCATATGCTGGATGGTTT	63				
QY	979 TACCTCAACTTTGTTTACTGAGAAGTAGAAACCCTGAAAAAACCATTCCTCCCTTGCATA	1038				

[illegible]

QY 378 CTCATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTTCATCTCTCTCTTAAGGGCGTGCT 437
Db 309 CTCATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTTCATCTCTCTCTTAAGGGCGTGCT 368
QY 438 CCAGAACACGGGCGGCGTGGGCGATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTC 497
Db 369 CCAGAACACGGGCGGCGTGGGCGATGCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTC 428
QY 498 ACTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACACTATAAAGAAATCTGGAGGTCA 488
Db 429 ACTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACACTATAAAGAAATCTGGAGGTCA 557
QY 558 TTACACATATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGA 617
Db 489 TTACACATATA-TTGGAGCTTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGA 547
QY 618 ACTCCTCATATAACGCCCTGCAGTACTGCTGTGATATCCCTGGCATTTGGACGCTACAT 677
Db 548 ACTCCTCATATAACGCCCTGCAGTACTGCTGTGATATCCCTGGCATTTGGACGCTACAT 607
QY 678 TCTGGAACCATTTTATTCAATGTGAATCCCTGAACTTGGCATCAAGCTCATTACAGC 737
Db 608 TCTGGAACCATTTTATTCAATGTGAATCCCTGAACTTGGCATCAAGCTCATTACAGC- 666
QY 738 TGTGGGCAATAACTGTAGTGGTCTCTAAATAGCATGAGTGTGAG 782
Db 667 TGTGGGCAATAACTGTAGTGGTCTCTAAATAGCATGAGTGTGAG 711

RESULT 6
LOCUS BG284503
DEFINITION 602408645F1 NIH_MGC_91 Homo sapiens cDNA clone linear EST 21-FEB-2001
mRNA sequence.
ACCESSION BG284503
VERSION BG284503.1 GI:13035516
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10463 row: i column: 11
High quality sequence stop: 795.
Location/Qualifiers
1..936
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4537810"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source
BASE COUNT 248 a 202 c 251 g 235 t
ORIGIN
Query Match 18.4%; Score 411.2; DB 12; Length 936;

Best Local Similarity 87.9%; Pred. No. 9.1e-69;
Matches 494; Conservative 0; Mismatches 13; Indels 55; Gaps 2;
QY 182 GAAAGCCTGTGTGTCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAAACGGGA 241
Db 278 GAAAGCCTGTGTGTCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAAACGGGA 337
QY 242 GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGACGAGCGCTTTTCAGGAAGAGACG 301
Db 338 GGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGACGAGCGCTTTTCAGGAAGAGACG 375
QY 302 CCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
Db 376 -----GAGAAAGTGCAGCTGAAGAGGAAAGTCA 403
QY 362 CTTTACTGAGGGGAGTCTCCATTATTCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCACT 421
Db 404 CTTTACTGAGGGGAGTCTCCATTATTCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCACT 463
QY 422 CTCCTAAGGGCGTGTCTCCAGAACACACGGGACGCGTGGGCATGTCTCTGACCATCTGGACGG 481
Db 464 CTCCTAAGGGCGTGTCTCCAGAACACACGGGACGCGTGGGCATGTCTCTGACCATCTGGACGG 523
QY 482 TGTGTGGGGTCTCTGTCACCTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACTATTA 541
Db 524 TGTGTGGGGTCTCTGTCACCTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACTATTA 583
QY 542 AGAAATCTGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTG 601
Db 584 AGAAATCTGAGGTCATTACACATATATCTCGGAAGTCTTTGGTCCATTACCAGCTTTTG 643
QY 602 TACGAGTCTGGGTGGAACCTCCTCATAATAACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661
Db 644 TACGAGTCTGGGTGGAACCTCCTCATAATAACGCCCTGCAGCTACTGCTGTGATATCCCTGG 703
QY 662 CATTGGACGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAACTTGGGA 721
Db 704 CATTGGACGCTACATTTCTGGAACCA-TTGTATTCAATGTGGAATCCCGAACTGGGA 762
QY 722 TCAAGTCTATTACAGCTGTGGG 743
Db 763 TCAAGTCTATTACAGCTGTGGG 784

RESULT 7
LOCUS W00842
DEFINITION Yz96dl2.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone linear EST 18-APR-1996
IMAGE:290903 5', mRNA sequence.
ACCESSION W00842
VERSION W00842.1 GI:1272822
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 371)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfsing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: ETPRimer
High quality sequence stop: 339.

Db 63 TTTTACTTCATTTTCTGAAAGTCTAGAGAATTACAACTTTGGTGATAACAAAA 10

RESULT 5
BG326527 910 bp mRNA linear EST 27-FEB-2001
LOCUS 602425373F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562994 5',
DEFINITION mRNA sequence.
ACCESSION BG326527
VERSION BG326527.1 GI:13132964
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 910)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1276 row: b column: 19
High quality sequence stop: 706.

FEATURES
source
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4562994"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 237 a 178 c 249 g 246 t
ORIGIN
Query Match 19.9%; Score 444.6; DB 12; Length 910;
Best Local Similarity 81.6%; Pred. No. 3.8e-75;
Matches 575; Conservative 0; Mismatches 74; Indels 56; Gaps 3;
QY 78 AAAGAAAGAAAGAGAGAGAGAGAAATCCAGGCCCAATTTGGCATFAGATTTTATCATAT 137
Db 63 AAGCAGAGGAAGACATCGATCAGTAACACCAAGAGAGACACCAAGTTGAAAGTTTGTTC 122
QY 138 TCTGGATTTTGGATTCCTTTTCTCATCTGATTCAGGAAAGCCTGTGTGTC 197
Db 123 TTTCCCTCTGTTTATTTTCCCTCCCTGTGTCTACTATGGTCAGAAAGCCTGTGTGTC 182
QY 198 CACCATCTCCAAAGGAGGTTTACCTGCAGGGAAATGTTAACGGAGGCTGCCCTGGG 257
Db 183 CACCATCTCCAAAGGAGGTTTACCTGCAGGGAAATGTTAACGGAGGCTGCCCTGGG 242
QY 258 CAACAAGGAGCCACCTGGCAGGACGCCTTTTCAGGAAGAGACGCCTTTTCAGGAAGAGA 317
Db 243 CAACAAGGAGCCACCTGGGCGAG----- 264
QY 318 CGCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAGTCACTTTACTGAGGGAGT 377
Db 265 -----GAGAAAGTGCAGCTGAAGAGGAAGTCACTTTACTGAGGGAGT 308

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NCI_CGAP_Sub3 library is a subtracted library derived from
the NCI_CGAP_Sub1 library, which is a subtracted library
derived from BI. BI constitutes a mixture of 21
normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4
, NCI_CGAP_Pr22, NCI_CGAP_Kid5, NCI_CGAP_Co10,
NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Lei2,
NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clones 1322376-1323911, 1456008-1456775
, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342
, 3722-3725, 3776-3778 (IMAGE Clones 1323912-1325831,
1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1
LLAM 3575-3582, 3851-3854 (IMAGE Clones 1414920-1417991,
1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 (IMAGE Clones 1257096-1258631,
1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones
985608-986759, 1101192-1101959, 1217928-1220615);
NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
Clones 1057416-1061255, 1144584-1145351). Subtraction
was performed as previously described [Bonaldo, Lennon &
Soares (1996): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.
TAG_LIB=NCI_CGAP_Co10
TAG_TISSUE=colon
TAG_SEQ=AAACG"

BASE COUNT 150 a 98 g 137 t
ORIGIN
Query Match 21.1%; Score 472.4; DB 10; Length 483;
Best Local Similarity 99.8%; Pred. No. 2.4e-80;
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1370 TCTCTGGAGACCTCGACAGTCTTGTGAATTTCCAGTTTGGCCAGGTGGCTTTTATTG 1429
Db 483 TCTCTGGAGACCTCGACAGTCTTGTGAATTTCCAGTTTGGCCAGGTGGCTTTTATTG 424
QY 1430 GGCTGGCAGTGTGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTT 1489
Db 423 GGCTGGCAGTGTGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTT 364
QY 1490 TCAAGGTGCCACTGTTTCATCCAGCTTGTTCCTTCACATGCCTCTTCATGTTGCC 1549
Db 363 TCAAGGTGCCACTGTTTCATCCAGCTTGTTCCTTCACATGCCTCTTCATGTTGCC 304
QY 1550 TTTCCCTCTATTTCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAG 1609
Db 303 TTTCCCTCTATTTCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAG 244
QY 1610 TCCCTGCGTATTATCTCTTATATATGGGACAAGAAACCAGGTGGTTTAGAATAATGT 1669
Db 243 TCCCTGCGTATTATCTCTTATATATGGGACAAGAAACCAGGTGGTTTAGAATAATGT 184
QY 1670 CAGAGAAATAACCAACACATTACAAATAATACTGGAAGTTGTACCAGAAGAGATAAGT 1729
Db 183 CAGAGAAATAACCAACACATTACAAATAATACTGGAAGTTGTACCAGAAGAGATAAGT 124
QY 1730 TATGAACATAATGGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGAT 1789
Db 123 TATGAACATAATGGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGAT 64
QY 1790 TTTTACTTCAATTTCTGAAGTCTAGAGAATTACAACCTTTGGTGATAACAAAA 1843

Accession BG490131
Version BG490131.1 GI:13451641
Keywords EST.
Source human.
Organism Homo sapiens
Reference NIH-MGC http://mgi.nci.nih.gov/
Authors National Institutes of Health, Mammalian Gene Collection (MGC)
Title Unpublished (1999)
Journal Contact: Robert Strausberg, Ph.D.
Comment Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTP/DTP/Gazdar
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
Plate: LCM1398 row: c column: 14
High quality sequence stop: 726.
Location/Qualifiers 1..919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4637509"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
Base Count 253 a 195 c 199 g 272 t
Origin
Query Match 21.1%; Score 473; DB 12; Length 919;
Best Local Similarity 76.2%; Pred. No. 1.4e-80;
Matches 728; Conservative 0; Mismatches 5; Indels 222; Gaps 3;
QY 829 GCAATTCGATAATTATAGTCCCTGGA-GTTATGCAGCTAATTAAGGTCACAGCAGAA 887
Db 2 GCAATTCGATAATTATAGTCCCTGGAGGTTATGCGAGCTAATTAAGGTCACAGCAGAA 61
QY 888 CTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTA 947
Db 62 CTTTAAAGACGCCCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTA 121
QY 948 TTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGA 1007
Db 122 TTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGA 181
QY 1008 AAACCCCTGAAAAAACCAATCCCTTGCAATATATATATCCATGGCCATGTACACATTGG 1067
Db 182 AAACCCCTGAAAA
QY 1068 CTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAAATGCTGAGGAGCTGCTGCTTC 1127
Db 194
QY 1128 AAATGCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTTCC 1187
Db 194
QY 1188 GATCTTTGTCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAG 1247
Db 238 GATCTTTGTCCTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCA- 296

mRNA sequence.
QY 1248 GTTATTCTATGTTGCGTCTCGAGAGGGTCACTTCCAGAAAATCCTCTCCATGATTCATGT 1307
Db 297
QY 1308 CCGCAAGCACACTCTCTACAGAGCTGTTATTGTTTGGCACCCTTTTGACAATGATAATGCT 1367
Db 297
QY 1368 CTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTAT 1427
Db 321 CTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTAT 380
QY 1428 TGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCC 1487
Db 381 TGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCC 440
QY 1488 TTTCAAGGTGCCACTGTTTCCAGCTTTGTTTTCCTTCACATGCCCTCTTTCATGGTTGC 1547
Db 441 TTTCAAGGTGCCACTGTTTCCAGCTTTGTTTTCCTTCACATGCCCTCTTTCATGGTTGC 500
QY 1548 CCTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTCTCATCTCTGACTGG 1607
Db 501 CCTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTCTCATCTCTGACTGG 560
QY 1608 AGTCCCTGCGTATTATCTTTTATATATATGGGACAAAGAAACCCAGGTGGTTTGAATAAT 1667
Db 561 AGTCCCTGCGTATTATCTTTTATATATATGGGACAAAGAAACCCAGGTGGTTTGAATAAT 620
QY 1668 GTCAGAGAAATAACCCAGAACATTTACAAATAATCTGGAAGTTGTACCAGAGAGATAA 1727
Db 621 GTCAGAGAAATAACCCAGAACATTTACAAATAATCTGGAAGTTGTACCAGAGAGATAA 680
QY 1728 GTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAA 1782
Db 681 GTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAA 735
RESULT 4
AW205572/c
LOCUS
DEFINITION UI-H-B11-adt-h-02-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
IMAGE:2718171 3', mRNA sequence.
ACCESSION AW205572
VERSION AW205572.1 GI:6505046
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 483)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=Yes.
Location/Qualifiers 1..483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2718171"
/clone_lib="NCI_CGAP_Sub3"
/lab_host="DH10B (Life Technologies)"
FEATURES
source

REFERENCE

AUTHORS Xu, H., Severson, T. and Acott, T.S.

TITLE Expression Profile and Genome Location of cDNA Clones from an Infant Human Trabecular Meshwork Library

BG490131
LOCUS
BG490131
919 bp
mRNA
linear EST 27-MAR-2001
Homo sapiens CDNA clone IMAGE:4637509 5'

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 05:48:49 ; Search time 3548 Seconds
(without alignments)
10220.320 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 2239

Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	512.8	22.9	516	9	AL702210
2	508.4	22.7	511	14	BM887840
3	473	21.1	919	12	BM887840 TMT023 Hu
4	472.4	21.1	483	10	AW205572
5	444.6	19.9	910	12	AW205572 UI-H-B11-
6	411.2	18.4	936	12	BM887840

7	363	16.2	371	14	W00842
8	359.8	16.1	497	14	BQ370588
9	326.2	14.6	1072	12	BG388107
10	311.6	13.9	1962	11	BC022457
11	250	11.2	403	13	BJ075046
12	243.2	10.9	872	14	BQ430089
13	228.4	10.2	945	14	BQ682698
14	225.4	10.1	524	12	BF286052
15	222.6	9.9	924	14	BQ684696
16	221.8	9.9	912	14	BQ894656
17	220.2	9.8	1182	11	AK005282
18	213.2	9.5	1068	14	BM920139
19	212.4	9.5	952	14	BQ219976
20	207	9.2	952	14	BQ680741
21	204.2	9.1	557	17	AQ002318
22	202	9.0	959	9	AL553189
23	201.8	9.0	876	14	BQ677699
24	201	9.0	1123	13	BM460946
25	200	8.9	682	9	AJ392529
26	200	8.9	702	9	AL635772
27	198.2	8.9	831	13	BI181265
28	197.2	8.8	991	13	BM556677
29	193.4	8.6	945	14	BQ682351
30	191.8	8.6	940	14	BQ678193
31	189.4	8.5	915	14	BQ681717
32	187.2	8.4	649	9	AL635943
33	186	8.3	530	9	AI102370
34	184	8.2	637	9	AL586948
35	182.8	8.2	581	12	BF717670
36	181.6	8.1	1155	13	BM474488
37	180.6	8.1	889	9	AL522431
38	180	8.0	949	14	BQ678681
39	179.8	8.0	927	14	BQ683707
40	179.4	8.0	1054	14	BM919004
41	177.4	7.9	706	12	BG761990
42	173.4	7.7	953	14	BQ683115
43	172.8	7.7	952	9	AL560263
44	172	7.7	551	14	BM962577
45	171.8	7.7	701	14	BM963289

ALIGNMENTS

RESULT 1
AL702210
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL702210
DKFZp686C09156_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
516 bp mRNA linear EST 22-MAR-2002
DKFZp686C09156 5', mRNA sequence.
AL702210
AL702210.1 GI:19685565
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 516)
Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. and
,S.
EST (Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. and
Wiemann,S.)
Unpublished (1999)
Contact: Bloecker H
MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp686C09156) is available at the RZPD in Berlin.


```
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
PS Disclosure; SEQ ID NO 23451; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 123 BP; 32 A; 31 C; 37 G; 23 T; 0 other;

Query Match 4.5%; Score 100; DB 22; Length 123;
Best Local Similarity 100.0%; Pred. No. 7.3e-32;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 GAAAGCCTGTTGTCCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAACGGGA 241
Db |||||
QY 14 GAAAGCCTGTTGTGTCCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAACGGGA 73
Db |||||
QY 242 GGCTGCCTTCCCTGGGCAACAAGGAGGCCACCTGGGCAGGA 281
Db |||||
QY 74 GGCTGCCTTCCCTGGGCAACAAGGAGGCCACCTGGGCAGGA 113
Db |||||

RESULT 15
ABN35211
ID ABN35211 standard; DNA; 60 BP.
XX
AC ABN35211;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:7959.
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
```


us=09-667-170a-440.oli.rng

Thu Apr 17 09:48:28 2003

```
QY 527 TGGGAACAACATAAAGAAATCTGGAGGTCAATACACATATATTTGG 574
      |||||||
Db 636 TGGGAACAACATAAAGAAATCTGGAGGTCAATACACATATATTTGG 683

RESULT 13
ABL37718
ID ABL37718 standard; cDNA; 575 BP.
XX
AC ABL37718;
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polynucleotide SEQ ID NO:1307.
XX
DE Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200196388-A2.
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18557.
XX
PR 09-JUN-2000; 2000US-210899P.
PR 20-FEB-2001; 2001US-270216P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Harlocker SL, Secríst H;
XX
PI WPI; 2002-114514/15.
DR
XX Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient
XX
PS Claim 1; SEQ ID 1307; 105pp; English.
XX
CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour.
XX
SQ Sequence 575 BP; 151 A; 122 C; 153 G; 141 T; 8 other;

Query Match 4.8%; Score 107; DB 24; Length 575;
Best Local Similarity 100.0%; Pred. No. 7.9e-35;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
      |||||||
Db 354 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 413

QY 394 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGGTGCTCCA 440
      |||||||
Db 414 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCCTAAGGGCGGTGCTCCA 460

RESULT 14
AAK68639
ID AAK68639 standard; DNA; 123 BP.
XX
AC AAK68639;
XX
```

```
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23451.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
```

AAZ16609	standard; cDNA; 800 bp.
AAZ16609;	
12-OCT-1999	(first entry)
Human gene expression product	cDNA sequence SEQ ID NO:4079.
Human; gene; gene expression product;	detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss
Homo sapiens.	
WO9938972-A2.	
05-AUG-1999.	
28-JAN-1999;	99WO-US01619.
03-APR-1998;	98US-0080666.
28-JAN-1998;	98US-0072910.
24-FEB-1998;	98US-0075954.
31-MAR-1998;	98US-0080114.
03-APR-1998;	98US-0080515.
(CHIR) CHIRON CORP.	
(HYSE-) HYSEQ INC.	
Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;	
Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;	
Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;	
Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;	
Stache-Crain B, Sudduth-Klinger J, Williams LT;	
WPI; 1999-494092/41.	
Novel human genes and their expression products which are	
differentially expressed in different cell types	
Claim 1; Page 1934; 2479pp; English.	

SQ Sequence 800 BP; 199 A; 166 C; 202 G; 211 T; 22 other;
 Query Match 4.8%; Score 108; DB 20; Length 800;
 Best Local Similarity 100.0%; Pred. No. 2.9e-35;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	467	TGACCATCTGGACGGTGTGGGGTCCTGTCACTATTGGAGCTTTGTCTTATGCTGAAT	526
Db	576	TGACCATCTGGACGGTGTGGGGTCCTGTCACTATTGGAGCTTTGTCTTATGCTGAAT	635

XX WPI; 2000-126369/11.
XX Polynucleotide library used to determine cancerous states of mammalian
XX cells -
XX Claim 1; Page 392; 1097pp; English.
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
XX libraries constructed from human colon cancer cell lines. The present
XX invention also describes a method of detecting differentially expressed
XX genes correlated with a cancerous state of a mammalian cell, comprising
XX detecting at least one differentially expressed gene product in a test
XX sample derived from a cell suspected of being cancerous, where detection
XX of the differentially expressed gene product is correlated with a
XX cancerous state of the cell from which the test sample was derived.
XX The polynucleotides sequences can be used in a method for detecting
XX differentially expressed genes correlated with a cancerous state of a
XX mammalian cell. The polynucleotides can also be used as probes for
XX detecting and mapping related genes. They can be used in diagnosis and
XX prognosis of diseases and disorders (e.g. identification of
XX pre-metastatic or metastatic cancerous states, stages of cancer, or
XX responsiveness of cancer to therapy). This is particularly for breast
XX cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
XX negative breast cancer, lung cancer, and colon cancer.
XX Sequence 300 BP; 80 A; 65 C; 66 G; 89 T; 0 other;

Query Match 12.7%; Score 284; DB 21; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.6e-109;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 CTGGCATTTGGACGCTACATCTCTGGAACCAATTTTATTCAATGTGAATCCCTGAACCT 717
DB 17 CTGGCATTTGGACGCTACATCTCTGGAACCAATTTTATTCAATGTGAATCCCTGAACCT 76
QY 718 GCGATCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGATGGTCTTAATAGCATGAGT 777
DB 77 GCGATCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGATGGTCTTAATAGCATGAGT 136
QY 778 GTCAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTCGT 837
DB 137 GTCAGCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTCGT 196
QY 838 ATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGAC 897
DB 197 ATAATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGAC 256
QY 898 GCCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGC 941
DB 257 GCCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGC 300

RESULT 10
AAD23462
ID AAD23462 standard; cDNA; 337 BP.
XX
AC AAD23462;
XX
XX
DT 26-FEB-2002 (first entry)
XX
XX Human lung tumour-specific 20E10 5' cDNA.
XX
XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; 20E10; ss.
XX
XX Homo sapiens.
OS
XX WO200172295-A2.
PN
XX
XX 04-OCT-2001.
PD
XX
XX 28-MAR-2001; 2001WO-US09991.
PF

XX 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-234517P.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
XX WPI; 2001-639201/73.
XX
XX New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer -
PT
XX Claim 1; Page 334; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and
CC their corresponding cDNA molecules. Lung tumour-specific proteins and
CC their antigen-presenting cells are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, and for inhibiting
CC the development of cancer. The invention also relates to a composition
CC useful for stimulating an immune response, and for treating cancer. The
CC lung tumour specific oligonucleotide is useful in gene therapy and for
CC diagnosis, detection and treatment of lung cancer. The present sequence
CC is a cDNA encoding human lung tumour-specific protein.
XX
XX Sequence 337 BP; 103 A; 60 C; 93 G; 81 T; 0 other;

Query Match 12.6%; Score 281; DB 23; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.5e-107;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAGGTTGAAGTGACGAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAC 60
DB 4 GGAGGTTGAAGTGACGAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAC 63
QY 61 AGAATTAAGGAAAAAGAAAAAGAAAAAGAGAGAGAGAGAAATCCAGGCCAATTGTGG 120
DB 64 AGAATTAAGGAAAAAGAAAAAGAAAAAGAGAGAGAGAGAGAAATCCAGGCCAATTGTGG 123
QY 121 CATAGATTTTATCATATTCTGGATTTTGGATTTCTTTTCTCATCTGATTTCA 180
DB 124 CATAGATTTTATCATATTCTGGATTTTGGATTTCTTTTCTCATCTGATTTCA 183
QY 181 GGAAGGCTTGTGTGTCACCATCTCCAAAGAGAGGTTACCTGCAGGAAATGTTAACGGG 240
DB 184 GGAAGGCTTGTGTGTCACCATCTCCAAAGAGAGGTTACCTGCAGGAAATGTTAACGGG 243
QY 241 AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCAGGA 281
DB 244 AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCAGGA 284

RESULT 11
ABK27651/c
ID ABK27651 standard; cDNA; 596 BP.
XX
XX
AC ABK27651;
XX
XX
DT 09-APR-2002 (first entry)
XX
XX Human colon cancer expressed sequence tag, Seq ID no 88.
DE
XX
XX Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW expressed sequence tag.
XX
XX Homo sapiens.
OS
XX WO200196390-A2.
PN

PR	28-JAN-1998;	98US-0072910.	XX	AAZ14962;	
PR	24-FEB-1998;	98US-0075954.	AC		
PR	31-MAR-1998;	98US-0080114.	XX	12-OCT-1999 (first entry)	
PR	03-APR-1998;	98US-0080515.	DT		
XX			XX	Human gene expression product cDNA sequence SEQ ID NO:2431.	
PA	(CHIR) CHIRON CORP.		DE		
PA	(HYSE-) HYSEQ INC.		XX		
XX	Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;		KW	Human; gene; gene expression product; diagnosis; therapy; probe;	
PI	Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;		KW	detection; mapping; tissue typing; profiling; forensic; cancer;	
PI	Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;		XX	genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.	
PI	Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;		OS	Homo sapiens.	
PI	Stache-Crain B, Sudduth-Klinger J, Williams LT;		XX		
XX			PN	WO9938972-A2.	
DR	WPI; 1999-494092/41.		XX		
XX			PD	05-AUG-1999.	
XX			XX		
PT	Novel human genes and their expression products which are		XX	28-JAN-1999; 99WO-US01619.	
PT	differentially expressed in different cell types		PF		
PT			XX		
XX	Claim 1; Page 1211; 2479pp; English.		XX		
XX			PR	03-APR-1998; 98US-0080666.	
XX			PR	28-JAN-1998; 98US-0072910.	
CC	The present invention describes a library of human polynucleotides		PR	24-FEB-1998; 98US-0075954.	
CC	comprising the sequences given in AAZ12532 to AAZ17779. Also described is		PR	31-MAR-1998; 98US-0080114.	
CC	a method of detecting differentially expressed genes correlated with the		PR	03-APR-1998; 98US-0080515.	
CC	cancerous state of a mammalian cell, comprising detecting at least one		XX		
CC	differentially expressed gene product in a test sample from a cell		PA	(CHIR) CHIRON CORP.	
CC	suspected of being cancerous, where the gene product is encoded by one		PA	(HYSE-) HYSEQ INC.	
CC	of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The		XX	Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;	
CC	polynucleotides can be used as a source of primers and probes, which can		PI	Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;	
CC	be used for a variety of purpose, e.g. detection of expression levels,		PI	Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;	
CC	mapping, tissue typing or profiling, forensics, genetic analysis and		PI	Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;	
CC	detection of polymorphisms. Polypeptides encoded by the polynucleotides		PI	Stache-Crain B, Sudduth-Klinger J, Williams LT;	
CC	can be used for raising antibodies for experimental, diagnostic and		XX		
CC	therapeutic purposes. The polynucleotides may also be used to construct		DR	WPI; 1999-494092/41.	
CC	arrays for diagnostics (which may be used to determine function of an		XX		
CC	encoded protein); and to detect differences in expression levels between		XX		
CC	two cells (e.g. to identify abnormal or diseased tissue in a human, to		PT	Novel human genes and their expression products which are	
CC	identify a genetic predisposition or susceptibility to a disease such as		PT	differentially expressed in different cell types	
CC	cancer). The polynucleotides of the invention are especially used in the		XX	Claim 1; Page 1211; 2479pp; English.	
CC	diagnosis, prognosis and management of colorectal cancer, breast cancer,		XX		
CC	and lung cancer. The polynucleotides can also be used to screen for		XX		
XX	peptide analogues and antagonists.		XX		
SQ	Sequence 300 BP; 58 A; 76 C; 67 G; 99 T; 0 other;		XX		
	Query Match 13.4%; Score 300; DB 20; Length 300;		XX		
	Best Local Similarity 100.0%; Pred. No. 1.8e-115;		XX		
	Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		XX		
QY	1067 GCTATGCTGACAAATGTGGCCCTACTTTACGACCATTAATGCTGAGGAGCTGCTTTT 1126		XX		
Db	1 GCTATGCTGACAAATGTGGCCCTACTTTACGACCATTAATGCTGAGGAGCTGCTTTT 60		XX		
QY	1127 CAAATGCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTC 1186		XX		
Db	61 CAAATGCAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTC 120		XX		
QY	1187 CGATCTTTGTTGCCCTCTCCCTGCTTGGCTCCATGAACGGTGGTGTGCTGTCTCCA 1246		XX		
Db	121 CGATCTTTGTTGCCCTCTCCCTGCTTGGCTCCATGAACGGTGGTGTGCTGTCTCCA 180		XX		
QY	1247 GGTATTCTATGTTGCGTCTCGAGGGGTCACTTCCAGAAATCCTCTCCATGATTCATG 1306		XX		
Db	181 GGTATTCTATGTTGCGTCTCGAGGGGTCACTTCCAGAAATCCTCTCCATGATTCATG 240		XX		
QY	1307 TCCGCAAGCAGCACTCCTCTACCACTGTTTATTGTTTGCACCCCTTTGACAAATGATAATGC 1366		XX		
Db	241 TCCGCAAGCAGCACTCCTCTACCACTGTTTATTGTTTGCACCCCTTTGACAAATGATAATGC 300		XX		
RESULT 7			XX		
AAZ14962			XX		
ID	AAZ14962 standard; cDNA; 300 BP.		XX		

Query Match 13.4%; Score 300; DB 20; Length 300;

Best Local Similarity 100.0%; Pred. No. 1.8e-115;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 300 BP; 85 A; 60 C; 63 G; 92 T; 0 other;

Query Match 13.4%; Score 300; DB 20; Length 300;

Best Local Similarity 100.0%; Pred. No. 1.8e-115;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 839 TAATTATAGTCCCTGGAGTTATGTCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGACG 898

Db 1 TAATTATAGTCCCTGGAGTTATGTCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGACG 60

QY	1834	ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTCT	1893
Db	1668	ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTCT	1727
QY	1894	AAGAAATTTAGTTATAAAGTCTATGTAGTTATAGAAAAGTGAATATGCAGTTATTCATGAG	1953
Db	1728	AAGAAATTTAGTTATAAAGTCTATGTAGTTATAGAAAAGTGAATATGCAGTTATTCATGAG	1787
QY	1954	TCGCACAATTCCTTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAAGACTAGACAA	2013
Db	1788	TCGCACAATTCCTTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAAGACTAGACAA	1847
QY	2014	TTACTATGTGGTCATTCCTACAACATATGTTAGCACGGCAAGAACCTTCAAAATTTGAAG	2073
Db	1848	TTACTATGTGGTCATTCCTACAACATATGTTAGCACGGCAAGAACCTTCAAAATTTGAAG	1907
QY	2074	ACTGAGATTTTCTGTATATATGGGTTTGG	2103
Db	1908	ACTGAGATTTTCTGTATATATGGGTTTGG	1937

RESULT 5
AAZ16528
ID AAZ16528 standard; cDNA; 772 BP.
.....

AC AAZ16528;

DT 12-OCT-1999 (first entry)

Human gene expression product cDNA sequence SEQ ID NO:3998.

Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss. XX OS Homo sapiens.

PN WO9938972-A2.

XX
PD 05-AUG-1999.

XX
PF 28-JAN-1999; 99WO-US01619.

PR 03-APR-1998; 98US-0080666.

PR 28-JAN-1998; 98US-0072910;

PR 24-FEB-1998; 98US-0075954;

31-MAR-1998; 98US-0080114.

03-APR-1998; 98US-0080515;

PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.

Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
Stache-Crain B, Sudduth-Klinger J, Williams LT;

WPI; 1999-494092/41.

Novel human genes and their expression products which are differentially expressed in different cell types

Claim 1; Page 1897; 2479pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ17779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The

polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for peptide analogues and antagonists.

SQ Sequence 772 BP; 175 A; 183 C; 158 G; 248 T; 8 other;

Query Match	18.5%;	Score 414;	DB 20;	Length 772;
Best Local Similarity	100.0%;	Pred. No. 3.7e-163;		
Matches 414; Conservative	0;	Mismatches	0;	Indels

Qy 1099 ACCATTAAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGG 1158
 |||||
 Db 90 ACCATTAAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGG 149

QY	1159	CTACTGGGAAATTTCTCATTAGCAGTTC CGATCTTGTGGCCCTCTCTCTGCTTTGGCTCC	1218
Db	150	CTACTGGGAAATTTCTCATTAGCAGTTC CGATCTTGTGGCCCTCTCTCTGCTTTGGCTCC	209

[illegible]

QY	1279	CTTCCAGAAATCCTCTCCATGATTTCATGTCCGCAAGCACACTCCTCTACCAAGCTGTTATT	1338
D _b	270	CTTCCAGAAATCCTCTCCATGATTTCATGTCCGCAAGCACACTCTCTACCAAGCTGTTATT	329

QY	1339	GTTTGGACCCCTTGACAATGATAATGCCTCTCTCTGGAGACCTCGACAGCTCTTTGAAT	1398
D _b	330	GTTTGGACCCCTTGACAATGATAATGCCTCTCTCTGGAGACCTCGACAGCTCTTTGAAT	389

1399	TTCCCTCAGTTTGGCCAGGTGGCTTTTATTTGGCTGGCAGTGC	TGGCTGGGCTGATTATCTT	1458
QY			
390	TTCCCTCAGTTTGGCCAGGTGGCTTTTATTTGGCTGGCAGTGC	TGGCTGGGCTGATTATCTT	449
Db			

	1459	CGATACAAATGCCCCAGATATGCATCGTCTTTCAAGGTGCCACTGTTTCATCCCCA	1512
QY			
	450	CGATACAAATGCCCCAGATATGCATCGTCTTTCAAGGTGCCACTGTTTCATCCCCA	503
Ddb			

RESULT 6
AAZ14980

IID AAZ14980 standard; cDNA; 300 BP.

AAZ14980;

12-OCT-1999 (first entry)

Human gene expression product cDNA sequence SEQ ID NO:2449.

Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

Homo sapiens.

WO9938972-A2

05-AUG-1999.

28-JAN-1999; 99WO-US01619.

R 03-APR-1998; 98US-0080666.

Thu Apr 17 09:48:28 2003

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.
XX Homo sapiens.
OS WO200154477-A2.
XX 02-AUG-2001.
XX 25-JAN-2001; 2001WO-US02687.
XX 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX WPI; 2001-476164/51.
DR P-PSDB; AAM23914.
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX Claim 1; Page 484; 1275pp; English.
XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention.
XX Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;
Query Match 76.8%; Score 1719; DB 22; Length 1958;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 334 GAGAAAGTGCAGCTGAGAGGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGSC 393
Db 168 GAGAAAGTGCAGCTGAGAGGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGSC 227
QY 394 ACCATCATTTGGAGCAGGAATCTTCACTCTCTCTAAGGGCGTGTCCAGAACACGGGCAGC 453
Db 228 ACCATCATTTGGAGCAGGAATCTTCACTCTCTCTAAGGGCGTGTCCAGAACACGGGCAGC 287
QY 454 GTGGGCATGTCTCTGACCATCTGGACGGTGTGGGGTCTGTCACTATTGGAGCTTTG 513
Db 288 GTGGGCATGTCTCTGACCATCTGGACGGTGTGGGGTCTGTCACTATTGGAGCTTTG 347
QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCAATTACACATATTTTG 573
Db 348 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCAATTACACATATTTTG 407
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGTGGAACTCCTCATATAATACGC 633
Db 408 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGTGGAACTCCTCATATAATACGC 467
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGCTACATTTCTGGAAACCATTTT 693
Db 468 CCTGCAGCTACTGCTGTGATATCCCTGGCATTGGACGCTACATTTCTGGAAACCATTTT 527
QY 694 ATTCAATGTGAATCCCTGAACCTTGCATCAAGCTCATTTACAGCTGTGGGCATTAAGTGA 753

Db 528 ATTCAATGTGAATCCCTGAACCTTGGGATCAAGCTCATTTACAGCTGTGGGCATTAAGTGA 587
QY 754 GTGATGGTCCCTAAATAGCATGAGTGTCAAGTGGAGCGCCCGGATCCAGATTTTCTTAACC 813
Db 588 GTGATGGTCCCTAAATAGCATGAGTGTCAAGTGGAGCGCCCGGATCCAGATTTTCTTAACC 647
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTTATAGTCCCTGGAGTTATGCAGCTAATATA 873
Db 648 TTTTGAAGCTCACAGCAATTTCTGATAATTTATAGTCCCTGGAGTTATGCAGCTAATATA 707
QY 874 GGTCAAAACGCAGAACCTTTAAAGACGCCCTTTTCAAGAGAGATTCAAGTATACGCGGTTG 933
Db 708 GGTCAAAACGCAGAACCTTTAAAGACGCCCTTTTCAAGAGAGATTCAAGTATACGCGGTTG 767
QY 934 CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 993
Db 768 CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 827
QY 994 ACTGAAGAAGTAGAAAACCTTGAAAAAACCATTTCCCTTGCATATGTATATCCATGGCC 1053
Db 828 ACTGAAGAAGTAGAAAACCTTGAAAAAACCATTTCCCTTGCATATGTATATCCATGGCC 887
QY 1054 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113
Db 888 ATTGTCAACCATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 947
QY 1114 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 948 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1007
QY 1174 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCTGCTTGGCTCCATGAACGGTGGTGTG 1233
Db 1008 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCTGCTTGGCTCCATGAACGGTGGTGTG 1067
QY 1234 TTTGCTGCTCCAGGTTATCTATGTTGGCTCTCGAGAGGGTCACTTCCAGAAATCCTC 1293
Db 1068 TTTGCTGCTCCAGGTTATCTATGTTGGCTCTCGAGAGGGTCACTTCCAGAAATCCTC 1127
QY 1294 TCCATGATTCATGTCGCAAGACACACTCTCTACAGCTGTTATTGTTTGCACCTTTG 1353
Db 1128 TCCATGATTCATGTCGCAAGACACACTCTCTACAGCTGTTATTGTTTGCACCTTTG 1187
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC 1413
Db 1188 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC 1247
QY 1414 AGGTGGCTTTTATTGGGCTGGCAGTGTGGCTGATTTATCTTCGATACAAATGCCCCA 1473
Db 1248 AGGTGGCTTTTATTGGGCTGGCAGTGTGGCTGATTTATCTTCGATACAAATGCCCCA 1307
QY 1474 GATATGATCGCTCTTCAAGGTGGCAGTGTTCATCCAGCTTTGTTTCTTCCATCATGC 1533
Db 1308 GATATGATCGCTCTTCAAGGTGGCAGTGTTCATCCAGCTTTGTTTCTTCCATCATGC 1367
QY 1534 CTCTTTCATGGTGGCTTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTC 1593
Db 1368 CTCTTTCATGGTGGCTTTCCCTCTATTCGGACCCATTTAGTACAGGGATTGGCTTCGTC 1427
QY 1594 ATCACTCTGAGTGGAGTCCCTGCGTATTTATCTTTATTTATGGGACAAAGAACCCAGG 1653
Db 1428 ATCACTCTGAGTGGAGTCCCTGCGTATTTATCTTTATTTATGGGACAAAGAACCCAGG 1487
QY 1654 TGGTTTAGAATAATGTCAGAGAAAATAACCAAGACATTTACAAATTAATCTGGAAGTTGTA 1713
Db 1488 TGGTTTAGAATAATGTCAGAGAAAATAACCAAGACATTTACAAATTAATCTGGAAGTTGTC 1547
QY 1714 CCAGAAAGAGATAAGTTATGAATAATGGAGTGTGAGATCTTGGCAATCTGCCCAAGGGGA 1773
Db 1548 CCAGAAAGAGATAAGTTATGAATAATGGAGTGTGAGATCTTGGCAATCTGCCCAAGGGGA 1607
QY 1774 GACACAAAATAGGGATTTTACTTCTATTTTCTGAAAGTCTAGAGAAATTTACAACTTTGGTG 1833
Db 1608 GACACAAAATAGGGATTTTACTTCTATTTTCTGAAAGTCTAGAGAAATTTACAACTTTGGTG 1667

CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;

Query Match 76.8%; Score 1719; DB 22; Length 1958;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	334	GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC	393
Db	168	GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGC	227
QY	394	ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTTAAGGGCGTCTCCAGAAACACGGGACG	453
Db	228	ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTTAAGGGCGTCTCCAGAAACACGGGACG	287
QY	454	GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTCACTATTGGAGCTTTG	513
Db	288	GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTGTCACTATTGGAGCTTTG	347
QY	514	TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGTCACTACACATATATTG	573
Db	348	TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGTCACTACACATATATTG	407
QY	574	GAAGTCTTTGGTCCATTACCAGCTTTGTACGAGTCTGGGTGGAACCTCTCCATAATACGC	633
Db	408	GAAGTCTTTGGTCCATTACCAGCTTTGTACGAGTCTGGGTGGAACCTCTCCATAATACGC	467
QY	634	CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT	693
Db	468	CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT	527
QY	694	ATTCAATGTGAATCCCTGAACTTGGATCAAGCTCAATACAGCTGTGGGCATAACTGTA	753
Db	528	ATTCAATGTGAATCCCTGAACTTGGATCAAGCTCAATACAGCTGTGGGCATAACTGTA	587
QY	754	GTGATGCTCCTAAATAGCATGAGTGTACGCTGGAGCGCCCGGATCCAGATTTCTTAACC	813
Db	588	GTGATGCTCCTAAATAGCATGAGTGTACGCTGGAGCGCCCGGATCCAGATTTCTTAACC	647
QY	814	TTTTGCAAGCTCACAGCAATCTGTATAATTATGATCCCTGGAGTATCAAGCTAATTA	873
Db	648	TTTTGCAAGCTCACAGCAATCTGTATAATTATGATCCCTGGAGTATCAAGCTAATTA	707
QY	874	GGTCAACGCAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTAAGCTA	933
Db	708	GGTCAACGCAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTAAGCTA	767
QY	934	CCACTGGCTTTTATATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT	993
Db	768	CCACTGGCTTTTATATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT	827
QY	994	ACTGAAGAAGTAGAAACCTTGAAACACCATTTCCCTTGCATATGATATGATATGATG	1053
Db	828	ACTGAAGAAGTAGAAACCTTGAAACACCATTTCCCTTGCATATGATATGATATGATG	887
QY	1054	ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTG	1113
Db	888	ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTACGACCATTAATGCTG	947
QY	1114	GAGCTGCTGCTTCAAATGCAAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC	1173
Db	948	GAGCTGCTGCTTCAAATGCAAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC	1007
QY	1174	TCATTAGCAGTCCGATCTTTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1233

Db	1008	TCATTAGCAGTCCGATCTTTGTTGGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1067
QY	1234	TTTGCTGTCTCCAGGTATCTATGTTGGTCTCGAGAGGGTCACTTCCAGAAATCTC	1293
Db	1068	TTTGCTGTCTCCAGGTATCTATGTTGGTCTCGAGAGGGTCACTTCCAGAAATCTC	1127
QY	1294	TCCATGATTTCATGTCGCGCAAGCACACTCTCTACCAGCTGTTATTTGCACCTTTG	1353
Db	1128	TCCATGATTTCATGTCGCGCAAGCACACTCTCTACCAGCTGTTATTTGCACCTTTG	1187
QY	1354	ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTGAAATTTCTCAGTTTGGC	1413
Db	1188	ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTGAAATTTCTCAGTTTGGC	1247
QY	1414	AGGTGGCTTTTATTGGGCTGGCAGTGTGGGCTGATTTATCTTATCTTATCTTATCTT	1473
Db	1248	AGGTGGCTTTTATTGGGCTGGCAGTGTGGGCTGATTTATCTTATCTTATCTTATCTT	1307
QY	1474	GATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATGTC	1533
Db	1308	GATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATGTC	1367
QY	1534	CTCTTCATGTTGGCTTTCCCTCTCTTTCGACCCATTTAGTACAGGGATTTGGCTTCGTC	1593
Db	1368	CTCTTCATGTTGGCTTTCCCTCTCTTTCGACCCATTTAGTACAGGGATTTGGCTTCGTC	1427
QY	1594	ATCCTCTGACTGGAGTCCCTGCGTATTATCTTATCTTATCTTATCTTATCTTATCTT	1653
Db	1428	ATCCTCTGACTGGAGTCCCTGCGTATTATCTTATCTTATCTTATCTTATCTTATCTT	1487
QY	1654	TGTTTAGAATAATGTCAGAGAAATAACCAAGAAATACAAATATCTTGAAGTTGTA	1713
Db	1488	TGTTTAGAATAATGTCAGAGAAATAACCAAGAAATACAAATATCTTGAAGTTGTC	1547
QY	1714	CCAGAAAGAGATAAGTATGAAGTATGAAGTCTGAGATCTTGGCAATCTGCCCCAAGGGA	1773
Db	1548	CCAGAAAGAGATAAGTATGAAGTATGAAGTCTGAGATCTTGGCAATCTGCCCCAAGGGA	1607
QY	1774	GACACAAATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAAATACAACTTTGTC	1833
Db	1608	GACACAAATAGGGATTTTACTTCAATTTCTGAAAGTCTAGAGAAATACAACTTTGTC	1667
QY	1834	ATAACAAAGAGTCAAGTATGATTTTATCTTATCTTATCTTATCTTATCTTATCTT	1893
Db	1668	ATAACAAAGAGTCAAGTATGATTTTATCTTATCTTATCTTATCTTATCTTATCTT	1727
QY	1894	AAGAAATTTAGTTATAAAGTCTATGATGATGATGATGATGATGATGATGATGATGAT	1953
Db	1728	AAGAAATTTAGTTATAAAGTCTATGATGATGATGATGATGATGATGATGATGATGAT	1787
QY	1954	TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTAGGGGTTAGGAGAAAGACTAGACA	2013
Db	1788	TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTAGGGGTTAGGAGAAAGACTAGACA	1847
QY	2014	TTACTATGTTGGTCAATCTCTCAACATATGTTTAGCAGCGCAAGAACTTCAAAATGAAG	2073
Db	1848	TTACTATGTTGGTCAATCTCTCAACATATGTTTAGCAGCGCAAGAACTTCAAAATGAAG	1907
QY	2074	ACTGAGATTTTCTGATATATGTTGTTT	2103
Db	1908	ACTGAGATTTTCTGATATATGTTGTTT	1937

RESULT 4
AAH98573
ID AAH98573 standard; cDNA; 1958 BP.
XX
AC AAH98573;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST-derived coding sequence SEQ ID NO: 430.

QY	1234	TTTGTGCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGGTCAACCTTCAGAAATCCTC	1293	XX	ABA09201;
Db	1178	TTTGTGCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGGTCAACCTTCAGAAATCCTC	1237	DT	11-JAN-2002 (first entry)
QY	1294	TCCATGATTCATGTCGCGAAGCACACTCCTCTACAGGTGTTATTGTTTGCACCCCTTG	1353	DE	Human cysteine/Glu transporter homologue cDNA, SEQ ID NO:977.
Db	1238	TCCATGATTCATGTCGCGAAGCACACTCCTCTACAGGTGTTATTGTTTGCACCCCTTG	1297	XX	Human; cytokine; cell proliferation; cell differentiation; growth factor;
QY	1354	ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGCC	1413	KW	haematopoiesis regulation; tissue growth; immunomodulator; activin;
Db	1298	ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGCC	1357	KW	inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
QY	1414	AGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA	1473	KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;
Db	1358	AGGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA	1417	KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
QY	1474	GATATGCATCGTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATCATGC	1533	KW	chronic inflammatory condition; proliferative retinopathy;
Db	1418	GATATGCATCGTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATCATGC	1477	KW	atherosclerosis; coronary heart disease; arterial ischaemia;
QY	1534	CTCTTCATGGTTGCCCTTTCCCTCTATTTCGACCCCATTTAGTACAGGGATTGGCTTCGTC	1593	KW	bone disorder; osteoporosis; vascular growth disorder;
Db	1478	CTCTTCATGGTTGCCCTTTCCCTCTATTTCGACCCCATTTAGTACAGGGATTGGCTTCGTC	1537	KW	tissue regeneration; wound healing; infection; immune disorder;
QY	1594	ATCACTCTGACTGGAGTCCCTCGTATTATCTCTTTATTATATGGACAAAGAACCCAGG	1653	KW	cell culture; drug screening; gene therapy; antiinflammatory;
Db	1538	ATCACTCTGACTGGAGTCCCTCGTATTATCTCTTTATTATATGGACAAAGAACCCAGG	1597	KW	antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
QY	1654	TGGTTTAGAATAATGTGAGAGAAAATAACCCAGAACATTACAAATAATACTGGAAGTTGTA	1713	KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
Db	1598	TGGTTTAGAATAATGTGAGAGAAAATAACCCAGAACATTACAAATAATACTGGAAGTTGTA	1657	KW	antifungal; vulnery; antiulcer; ss.
QY	1714	CCAGAAAGAGATAAGTTATGAACTAATGGACCTTGAGATCTTGGCAATCTGCCAAAGGGGA	1773	XX	Homo sapiens.
Db	1658	CCAGAAAGAGATAAGTTATGAACTAATGGACCTTGAGATCTTGGCAATCTGCCAAAGGGGA	1717	PI	WO200157188-A2.
QY	1774	GACACAAATAGGGATTTTACTTCATTTTCTGAAAGTCTAGAGAATTACAACTTTGGTG	1833	XX	09-AUG-2001.
Db	1718	GACACAAATAGGGATTTTACTTCATTTTCTGAAAGTCTAGAGAATTACAACTTTGGTG	1777	PT	05-FEB-2001; 2001WO-US03800.
QY	1834	ATAACAAAGAGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCT	1893	PT	03-FEB-2000; 2000US-0496914.
Db	1778	ATAACAAAGAGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCT	1837	PS	27-APR-2000; 2000US-0560875.
QY	1894	AAGAAATTTAGTTATACTCTATGTTAGTTATAGAAAGTGAATATGCAGTTATTTATGAG	1953	XX	(HYSE-) HYSEQ INC.
Db	1838	AAGAAATTTAGTTATACTCTATGTTAGTTATAGAAAGTGAATATGCAGTTATTTATGAG	1897	CC	Tang YT, Liu C, Drmanac RT;
QY	1954	TCGCACAATCTTGAGTCTCTGTATACCTACCTATTGGGGTTAGGAGAAAAGACTAGACAA	2013	CC	WPI; 2001-457740/49.
Db	1898	TCGCACAATCTTGAGTCTCTGTATACCTACCTATTGGGGTTAGGAGAAAAGACTAGACAA	1957	CC	P-PSDB; ABB11957.
QY	2014	TTACTATGGTGCAATCTCTACAAATATGTTAGCACGGGCAAGAACCTTCAAATTTGAAG	2073	CC	Human proteins and DNA encoding sequences useful for preventing,
Db	1958	TTACTATGGTGCAATCTCTACAAATATGTTAGCACGGGCAAGAACCTTCAAATTTGAAG	2017	CC	treating or ameliorating a medical condition in a mammalian subject
QY	2074	ACTGAGATTTTCTGATATATGGGTTTGTAAAGATGGTTTACACACTACAGATGTCT	2133	CC	e.g. arthritis and cancer
Db	2018	ACTGAGATTTTCTGATATATGGGTTTGTAAAGATGGTTTACACACTACAGATGTCT	2077	CC	Claim 1; Page 835; 1963pp; English.
QY	2134	ATACTGTGAAAAGTGTTCATTTCTGAAAAAAGCATACATCATGATTATGGCAAGAG	2193	CC	Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
Db	2078	ATACTGTGAAAAGTGTTCATTTCTGAAAAAAGCATACATCATGATTATGGCAAGAG	2137	CC	sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
QY	2194	GAGAGAA 2200		CC	invention also relates to vectors and recombinant host cells comprising a
Db	2138	GAGAGAA 2144		CC	nucleotide of the invention, methods of producing the novel polypeptides,
				CC	antibodies against the polypeptides, methods of detecting the nucleotides
				CC	or polypeptides in a sample, and methods of identifying compounds which
				CC	bind to polypeptides of the invention. Although novel, many of the
				CC	polypeptides of the invention have homology to known proteins, thereby
				CC	giving an insight into their probable biological activities, and hence
				CC	potential therapeutic applications. The polypeptides of the invention may
				CC	have various activities, including cytokine, cell proliferation or cell
				CC	differentiation activities; stem cell growth factor activity;
				CC	haematopoiesis regulatory activity; tissue growth activity;
				CC	immunomodulatory activity; activin- or inhibin-related activities;
				CC	chemotactic or chemokinetic activities; haemostatic, thrombotic or
				CC	thrombolytic activities; receptor or ligand activities; or may be
				CC	involved in oncogenesis, cancer cell proliferation or metastasis.
				CC	Depending on their biological activities, polypeptides and nucleotides of
				CC	the invention are useful for preventing, treating or ameliorating medical
				CC	conditions, e.g., by protein or gene therapy. Such conditions include
				CC	cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
				CC	disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
				CC	proliferative retinopathy, atherosclerosis, coronary heart disease,
				CC	arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
				CC	vascular growth. Polypeptides involved with tissue regeneration and
				CC	repair (or nucleic acids encoding them) may be used to promote wound
				CC	healing (e.g., of burns, incisions and ulcers), while those with
				CC	immunomodulatory activities may be used in the treatment of viral,

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 09:16:28 ; Search time 522 Seconds
(without alignments)
9659.433 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	100.0	2239	23	Human lung tumour-
2	1867	83.4	5981	23	Human lung tumour-
3	1719	76.8	1958	22	Human cystine/Glu
4	1719	76.8	1958	22	Human EST-derived
5	414	18.5	772	20	Human gene express
6	300	13.4	300	20	Human gene express
7	300	13.4	300	20	Human gene express
8	288	12.9	288	20	Human gene express
9	284	12.7	300	21	Human colon cancer

c	10	281	12.6	337	23	AAD23462	Human lung tumour-
	11	269	12.0	596	24	ABK27651	Human colon cancer
	12	108	4.8	800	20	AAZ16609	Human gene express
	13	107	4.8	575	24	ABL37718	Human colon tumour
	14	100	4.5	123	22	AAK68639	Human immune/haema
	15	60	2.7	60	24	ABN35211	Human spliced tran
	16	51	2.3	761	20	AAZ17094	Human gene express
	17	45	2.0	983	20	AAZ53014	Human prostate tum
	18	33	1.5	136	22	AAK82863	Human immune/haema
	19	33	1.5	136	22	AAK82864	Human immune/haema
	20	33	1.5	8948	22	AAK67209	Human immune/haema
	21	33	1.5	175737	24	ABK83571	Human cDNA differe
c	22	32	1.4	300	20	AAZ13802	Human gene express
	23	32	1.4	300	20	AAZ13190	Human gene express
	24	32	1.4	2440	24	ABL64408	Stomach cancer rel
c	25	32	1.4	3403	22	AAH18197	Human cDNA sequenc
c	26	32	1.4	8894	23	ABK42719	Genomic sequence #
	27	31	1.4	300	20	AAZ14169	Human gene express
	28	31	1.4	1443	22	AAD17386	Human bone marrow
	29	30	1.3	356	24	ABL83406	Human ovarian canc
c	30	30	1.3	569	22	AAH10296	Human cDNA clone (
c	31	30	1.3	804	22	AAH06560	Human cDNA clone (
c	32	30	1.3	2246	22	AAH16261	Human cDNA sequenc
	33	30	1.3	25012	22	ABA15431	Human nervous syst
	34	30	1.3	31804	22	AAK90698	Human digestive sy
	35	30	1.3	38894	22	AAK69848	Human immune/haema
c	36	29	1.3	482	24	ABL87814	Human ovarian canc
c	37	29	1.3	492	24	ABL82980	Human ovarian canc
c	38	29	1.3	501	24	ABL80020	Human immune syste
c	39	28	1.3	17674	24	ABL33345	Human prostate exp
	40	27	1.2	359	23	ABV01115	Human prostate exp
	41	27	1.2	360	23	ABV10284	Human prostate exp
	42	27	1.2	407	23	ABV31455	Human prostate exp
	43	27	1.2	407	23	ABV40424	Human prostate exp
	44	27	1.2	737	20	AAK02169	Human transcriptio
	45	27	1.2	3637	22	ABA20309	Human nervous syst

ALIGNMENTS

RESULT 1
AAD23460
ID AAD23460 standard; cDNA; 2239 BP.
XX
AC AAD23460;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human lung tumour-specific 19A4 cDNA.
XX
KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; 19A4; ss.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US09991.
XX
PR 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-234517P.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;

Thu Apr 17 09:48:35 2003

QY 1234 TTTGCTGTCTCCAGGTTATTTCTATGTTGCTCTCGAGAGGGTCCACCTTCCAGAAATCCTC 1293
Db 1270 TTTGCTTTCATCAAGGTTGTTCTCTGTTGGGCTCCCGGAGGGCCACCTACCGGACCTTCTG 1329
QY 1294 TCCATGATTCATGTCGCAAGCACACACTCTCTACAGCTGTTATTTGTTGACCCCTTTG 1353
Db 1330 TCCATGATCCACATGAGCGTTTACAGCTATCCCTGCTTACTGTTCAATGACCATG 1389
QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGCC 1413
Db 1390 GCACTCATCTACCTCATCTGAGGAGGATGTTTCCAGCTTATCAACTACTTCCAGCTTCCAGC 1449
QY 1414 AGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 1450 TACTGGTCTTCTGTTGGGCTGCTGTTGTTGGACAGCTCTACCTCCGCTGGAAGGAGCCC 1509
QY 1474 GATATGCATCGTCCCTTTCAAGGTGCCACTGTTTCATCCAGCTTTTCTTCCCTTCCATGC 1533
Db 1510 AAGCGCCCCCGGCTCTCAAGCTGAGCGTGTGTTTCCCATCGTGTCTGTCATATGCTCC 1569
QY 1534 CTCCTCATGGTGGCTTCCCTTCCCTCTATTCGAGCCCATTTAGTACAGGATTTGCTTCTGTC 1593
Db 1570 GTGTTCTGGTGATAGTGGCCCTCTTCACTGACACCAATTAATCCCTCATTTGGCATCGGG 1629
QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATT 1621
Db 1630 ATTGCCCTTTCTGGAGTCCCTTCTACT 1657

Search completed: April 16, 2003, 10:55:37
Job time : 241 secs

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGGAGTCTCCATATCATTTGGC 393
Db 370 GAAACTATGAGCTGAAGAGGAGATCTCCCTGCTGAATGGGTGAGCTGCTGGTGGG 429
QY 394 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTAAGGGGCTGCTCCAGAACACGGGCAGC 453
Db 430 AACATGATCGGCTCAGGGATCTTTGTCTCACCCAAAGGGTGTGTTACACACTGCTCC 489
QY 454 GTGGGCTGCTCTGACCATCTGGAGCGTGTGGGGTCTGCTCACTATTTGGAGCTTTG 513
Db 490 TATGGGATGTCACATGATTTGTGGGCAATGTTGGGCTCTCTCTGTTGTGGGTGCCCTT 549
QY 514 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATACACATATATTG 573
Db 550 TGTATGACAGCTGGGACCAACCATACCAAGTGGGAGCCAGCTACGCTTATATCTA 609
QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTAGAGTCTGGGTGGAACTCCTCATATAACGC 633
Db 610 GAGGCTTTGGGGCTTCATGCTTCCCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 669
QY 634 CCTGAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 693
Db 670 CCCACGGTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
QY 694 ATTCATGTAATCCCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753
Db 730 CCCAGTGTGATCCCCCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789
QY 754 GTGATGCTCTAAATAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 813
Db 790 CTGACATTTGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
QY 814 TTTTGAAGTCAAGCAAGCAATTTCTGATTAATATAGTCCCTGGAGTTATGACGCTAATTA 873
Db 850 TACGCCAAGTGTGAGGCTCATTTGCCATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
QY 874 GGTCAAAACGCAAGCACTTTAAAGCGCTTTTTCAGGAAGAGATTCAGATATTACGCGGTTG 933
Db 910 GGACACTCTGAGCACTTTTCAAGCGCTTTTGGAGGTTTCTCTGCTGCTGCTGCTGCTGCTG 969
QY 934 CCAGTGGCTTTTATGATGGAATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 993
Db 970 TCTCTTGGCTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
QY 994 ACTGAAGAAGTAGAAACCTTGAAACCAATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1053
Db 1030 ACAGAAGAATCAAAACCAACCAAGAAATTTGCCCTTGGCCATTTGGGATTTCTATGCCA 1089
QY 1054 ATTGTCACCATTTGGCTATGCTGCTGACAAATGTTGGCTACTTTACGACCATTAATGCTGAG 1113
Db 1090 ATTGTGACGCTCATCTACATCTGACCAATGTTGGCTATTTACAGAGTGTGAACATTTCA 1149
QY 1114 GAGTGTGCTGCTTTCAAAATGAGTGGCAGTGAACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 1150 GATGTCTTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
QY 1174 TCATTAGCAGTTCCTGATCTTTGTTGCCCTCTCTGCTGCTTGGCTCCATGAACGGTGGTGTG 1233
Db 1210 AGCTGGACCATCCCCATTTGCTGTTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269

Db 1328 GTGTTTCTGGTATAGTGGCCCTCTTCACTGACACCATTAATTCCTCATTTGGCATCGG 1387

QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATT 1621

Db 1388 ATTGCCCTTTCTGGAGTCCCTTTTCTACT 1415

RESULT 14

US-10-163-866-8

; Sequence 8, Application US/10163866

; Publication No. US20030027188A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/10/163,866

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: US 60/357,600

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 6295

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-163-866-8

Query Match 16.9%; Score 379.2; DB 9; Length 6295;

Best Local Similarity 55.9%; Pred. No. 4.1e-91;

Matches 720; Conservative 0; Mismatches 568; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393

Db 363 GAAACTATGCAGCTGAAGAGGAGATCTCCCTGCTGAATGGGTGAGCCTGGTGGGC 422

QY 394 ACCATCATTTGGAGCAGGAATCTTCACTCTCTCTTAAGGGCTGCTCCAGAACACGGGCAGC 453

Db 423 AACATGATCGGCTCAGGATCTTTGTCTCACCCAAAGGTGCTGGTACACACTGCCTCC 482

QY 454 GTGGGATGCTCTGACCATCTGACCGGTGTGGGGTCTGCTCACTATTGGAGCTTTG 513

Db 483 TATGGATGCTCACTGATTGTGGGCCATTGGTGGGCTCTTCTGTTGGGTGCCCCT 542

QY 514 TCTTATGCTGAATTGGGAACAACACTATAAAGAAATCTGGAGGTCAATTACACATATATTG 573

Db 543 TGTATGACAGAGCTGGGGACCACTACCAAGTCGGGAGCCAGCTAGCTATATTCTA 602

QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACAGTCTGGGTGGAACCTCCTCATAATACGC 633

Db 603 GAGGCTTTGGGGCTTCACTTGCCTTCACTCCGCTGTGGTCTCACTAGTTGTTGAG 662

QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCAATTTGGACGCTACATTTCTGGAACCATTTT 693

Db 663 CCACCGGTGAGGCCCATCATCGCCATCACCTTTGCCAACTACATCATCCAGCCGCTCTC 722

QY 694 ATTCAATGTGAATCCCTGAACCTTGGATCAAGCTCAATTACAGCTGTGGGCATCACTGTA 753

Db 723 CCCAGCTGTATCCCCCATACCTGGCCCTGCGCTCTCCTGGCTGCTGCTTGCATATGCTG 782

QY 754 GTGATGGTCTTAATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACC 813

Db 783 CTGACATTTGTGAAGTGTGCCTATGTCAAGTGGGCACACGTTGTCAGGACACGTTCACT 842

QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTATAGTCCCTGAGGTTATGCAGCTAATTA 873

Db 843 TAGCCCAAGGTCGTAGCGCTCATTTGCCATCATTTGTCATGGCCCTTGTAAACTGTGCCAG 902

QY 874 GGTCAAACGCAGAACTTTAAAGACGCCTTTTTCAGAAAGAGATTCAAGTATTACGGGTTG 933

Db 903 GGACACTCTGAGCACTTTTCAAGGACGCCTTTTGAGGGTCTCCTCGGACATGGGAAACCTC 962

QY 934 CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGTGGTTTACCTCAACTTTGTT 993

Db 963 TCTCTTGGCCTCTACTCTGCCCCTCTTCTTCTTACTCAGTTGGGACACCCCTTAATTTGTA 1022

QY 994 ACTGAAGAAGTAGAAACCCCTGAAAAAACCCATTCCTTGTCAATATATATATATATATAT 1053

Db 1023 ACAGAAGAAATCAAAAAACCCAGAAAGAAATTTGCCCTTGGCCATTTGGGATTTCTATGCCA 1082

QY 1054 ATTGTCACCATTTGGCTATGTGCTGACAAATGTGGCCTACTTTTACGACCATTAATGCTGAG 1113

Db 1083 ATTGTGACGCTCATCTACATCTGACCAATGTGGCCTATTACACAGTGTGAACATTTCA 1142

QY 1114 GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGCACCTTTTCTGAGCGGCTACTGGGAAATTT 1173

Db 1143 GATGCTCTTACAGTGTGCTGCTGGCTGTCCCTGCTTGGGGGCTCAATGATCCATC 1202

QY 1174 TCATTAGCAGTTCCGATCTTTGTTGGCCTCTCTCTACCACTGTTTACTGTTCAATGCAACCATG 1233

Db 1203 AGCTGGACCATCCCATTTGCTGCTGGCTGTCCCTGCTTGGGGGCTCAATGATCCATC 1262

QY 1234 TTTGCTGCTCCAGGTTATTCTATGTTGCGTCTCGAGAGGCTACCTTCCAGAAATCCTC 1293

Db 1263 TTTGCTTCAATCAAGGTTGTTCTGCTGGCTCCCGGAGGGCCACCTACCGGACCTTCTG 1322

QY 1294 TCCATGATTCATGTCGCAAGCAGCACACTCTCTACCACTGTTTACTGTTTGCACCTTTG 1353

Db 1323 TCCATGATTCACATTTGAGCGTTTACACCTATCCCTGCTTACTGTTCAATGCAACCATG 1382

QY 1354 ACAATGATAATGCTCTCTCTGAGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGC 1413

Db 1383 GCACATCATCTACCTCATCTGCTGGAGGATGTTTTCAGCTTATCAACTACTTCACTGCTT 1442

QY 1414 AGTGGCTTTTATTTGGGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473

Db 1443 TACTGGTTCTTCTGTTGGGCTGTCTGTTGTTGGAGAGCTCTACCTCCGCTGGAGGAGCCC 1502

QY 1474 GATATGATCGTCTCTTCAAGGTGCCACTGCTCATCCAGCTTTGTTTCTTCCATCATGC 1533

Db 1503 AAGCGGCCCCGCTCTCAAGCTGAGCGTGTGTTTCCCATCGTGTCTGCTATGCTTCC 1562

QY 1534 CTCTTCATGTTGGCTTTTCCCTCTATTTCGGACCCCATTTAGTACAGGATTTGGCTTCTG 1593

Db 1563 GTGTTTCTGTTGATAGTGGCCCTCTTCACTGACACCATTAATTCCTCATTTGGCATCGGG 1622

QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATT 1621

Db 1623 ATTGCCCTTTCTGGAGTCCCTTTTCTACT 1650

RESULT 15

US-10-163-866-9

; Sequence 9, Application US/10163866

; Publication No. US20030027188A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/10/163,866

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

Db 106 -----GAGAAAGTGCAGCTGAAGAGGAAAGTCA 133
QY 362 CTTTACTGAGGGAGCTCCATTATCATTTGACCATCATTTGGAGCAGGAATCTTCAATCT 421
Db 134 CTTTACTGAGGGAGCTCCATTATCATTTGACCATCATTTGGAGCAGGAATCTTCAATCT 193
QY 422 CTCCTAAGGGCGTGTCCAGAACACGGGCGGTGGGCATGTCTCTGACCATCTGGACGG 481
Db 194 CTCCTAAGGGCGTGTCCAGAACACGGGCGGTGGGCATGTCTCTGACCATCTGGACGG 253
QY 482 TGTGTGGGTCTCTGTCACTATTTGGAGCTTTTGTCTTATGTCTGAATTTGGGAACAATAA 541
Db 254 TGTGTGGGTCTCTGTCACTATTTGGAGCTTTTGTCTTATGTCTGAATTTGGGAACAATAA 313
QY 542 AGAAATCTGGAGGTCAATTACACATATATTTTGGAGTCTTTGGTCCATTACCAGCTTTTG 601
Db 314 AGAAATCTGGAGGTCAATTACACATATATTTTGGAGTCTTTGGTCCATTACCAGCTTTTG 373
QY 602 TACGAGTCTGGGTGGAACCTCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661
Db 374 TACGAGTCTGGGTGGAACCTCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 433
QY 662 CATTTGGAGCTACATTTCTGGAACCATTTTATTTCAATGTGAATCCCTGAACTTGCGA 721
Db 434 CATTTGGAGCTACATTTCTGGAACCATTTTATTTCAATGTGAATCCCTGAACTTGCGA 493
QY 722 TCAAGCTCATTTACAGCTGTGGGCATAA 748
Db 494 TCAAGCTCATTTACAGCTGTGGGCATAA 520

RESULT 13
US-10-163-866-10
; Sequence 10, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 1581
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-10

Query Match 16.9%; Score 379.2; DB 9; Length 1581;
Best Local Similarity 55.9%; Pred. No. 1.6e-91;
Matches 720; Conservative 0; Mismatches 568; Indels 0; Gaps 0;
QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
Db 128 GAAACTATGCAGCTGAAGAGGAGATCTCCTGCTGAATGGGGTCAGCCTGGTGGGC 187
QY 394 ACCATCATTTGGAGCAGGAATCTTCACTCTCTCCTAAGGGCGTGTCTCCAGAACACGGGAGC 453
Db 188 AACATGATCGGCTCAGGGATCTTTGTCTACCCAAAGGGTGTGCTGTGTACACACTGCCTCC 247
QY 454 GTGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACTATTTGGAGCTTTG 513

Db 248 TATGGGATGTCACTGATTTGTGGGCCATTTGGTGGGCTCTTCTCTGTTGGGTGCCCTT 307
QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGGTCAATTACACATATATTTG 573
Db 308 TGTATGCAGAGCTGGGAGCACCAACATCACCAAGTCGGGAGCCAGCTACGCTTATATCTA 367
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATACGC 633
Db 368 GAGGCTTTTGGGGCTTTCATTGGCTTTCATCCGCTGCTGCTGCTGCTGCTGCTGCTG 427
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTTACATTTCTGGAACCATTTT 693
Db 428 CCCACGGTTCAGGCCATCATCGCCATCACCTTTGCCAATCATCATCATCATCATCATC 487
QY 694 ATTCAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTTACAGCTGTGGGCATACCTGTA 753
Db 488 CCCAGCTGTATCCCCCATACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547
QY 754 GTGATGCTCCTAAATAGCATGAGTGTGAGCTGGAGCGGCCCGGATCCAGATTTTCTTAACC 813
Db 548 CTGACATTTGTGAACCTGTGCTATGTCAAGTGGGGCACACGCTGTGAGGACACGTTCACT 607
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATTTAGTCCCTGGAGTTATGTCAGCTAATTTAA 873
Db 608 TACGCCAAGTCTGAGCTCATTTGCCATCATTTGCTCATGCGCTTGTAACTGTGCCAG 667
QY 874 GGTCAACGCGAGCACTTTAAAGACGCCCTTTTCAAGAGAGATTTCAAGTATTACGCGTTG 933
Db 668 GGACACTCTGAGCACTTTTACGAGCGCTTTTGGAGTTCCTCTCTGGGACATGGGAAACCTC 727
QY 934 CCACTGGCTTTTATTTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGTT 993
Db 728 TCTCTTGGCTCTACTCTGCTCTTCTCTTACTCAGTTGGGACACCCCTTAATTTGTA 787
QY 994 ACTGAAGAGTAGAAAACCCCTGAAAACACCATTTCCCTTGCATATGTATATCCATGGCC 1053
Db 788 ACAGAAGAAATCAAAAACCCAGAAAGAAATTTGCCCTTGGCCATTTGGGATTTCTATGCCA 847
QY 1054 ATTGTCAACATGGCTATGTGCTGACAAATGTGGCTTACCTTACGACCATTTAATGCTGAG 1113
Db 848 ATTGTGACGCTCATCTACATCTCTGACCAATGTGGCTTATACACAGTGTGAACATTTCA 907
QY 1114 GAGCTGCTGCTTCAAAATGCAGTGGAGTGACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 908 GATGTCCTTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 967
QY 1174 TCATTAGCAGTTCGATCTTTGTTGCCCTCTCTCTGCTTGTGGCTTCCATGAACGCTGTG 1233
Db 968 AGCTGGACCATCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027
QY 1234 TTTGCTGTCTCCAGGTATTTCTATGTTGGTCTCGAGAGGGTCACTTCCAGAAATCCCTC 1293
Db 1028 TTTGCTTCATCAAGGTGTCTTCTGTTGGGCTCCCGGAGGGCCACCTACCGGACCTTCTG 1087
QY 1294 TCCATGATTCATGTCGCGAAGCACACTCTCTACAGCTGTATTGTTTGCACCCCTTTG 1353
Db 1088 TCCATGATTCACATTTGAGCGTTTACACCTATCCCTGCTTTACTGTTCAATTTGCACCATG 1147
QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCGCC 1413
Db 1148 GCACATCATCTACCTCATCTGAGGAGATGTTTTCAGCTTATCAACTACTTCACTTTCAGC 1207
QY 1414 AGGTGGCTTTTATTTGGGCTGGCAGTGTGCTGGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 1208 TACTGTTCTTCTGTTGGCTCTGCTGTTGTTGGACAGCTCTACCTCCGCTGGAAGGAGCCC 1267
QY 1474 GATATGCATGCTCTCTTCAAGGTGCCACTGTTTATCTCCAGCTTTGTTTCTTCCATGCTC 1533
Db 1268 AAGCGGCCCCGGCTCTCAAGCTGAGCGTGTGTTTCCCATCTGCTGTTCTGCTATGCTCC 1327
QY 1534 CTCTTTCATGTTGCCCTTTTCCCTCTATTTCCGACCCCTTATAGTACAGGGATTTGCTTCGTC 1593

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-53

Query Match 46.8%; Score 1047.4; DB 9; Length 1268;
Best Local Similarity 87.9%; Pred. No. 1.7e-271;
Matches 1238; Conservative 0; Mismatches 1; Indels 170; Gaps 2;

QY 182 GAAAGCCTGTGTGTCACCATCTCCAAAGGAGGTTTACCTGCAGGGAATGTTAACGGGA 241
Db |||||||
QY 242 GGCTGCCTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTTCAGGAAGAGAGC 301
Db |||||||
QY 302 CCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
Db |||||||
QY 362 CTTTACTGAGGGGAGTCTCCATATATCATTTGGCACCATCATTTGGAGCAGGAATCTTTCATCT 421
Db |||||||
QY 422 CTCTAAGGGCGTGTCTCAGAACACGCGGCGAGCGTGGGCATGTCTGTGACCATCTGGACGG 481
Db |||||||
QY 482 TGTGTGGGCTCCTGTCTACTATTTGGAGCTTTGTCTTATGCTGAATGGGAACAACATAAA 541
Db |||||||
QY 542 AGAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACACAGCTTTTG 601
Db |||||||
QY 602 TACGAGTCTGGGTGGAACCTCCTCATAATACGCCCTGCAGTACTGTGTGATATCCCTGG 661
Db |||||||
QY 662 CATTTGGACGTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACTTGC 721
Db |||||||
QY 722 TCAAGCTCATTACAGCTGTGGGCATACTGTAGTGGTCTTAAATAGCATGAGTGCA 781
Db |||||||
QY 782 GCTGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 841
Db |||||||
QY 842 TTATAGTCCCTGGAGTTATGACAGCTAATTAAGGTCAAACGCAGAACTTTAAAGACGCCT 901
Db |||||||
QY 902 TTTTCAAGGAGAGATTCAAGTATTACGGGTGGTCCACTGGCTTTTATTATGGAATGATG 961
Db |||||||
QY 962 CATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAA 1021
Db |||||||
QY 1022 CCATCCCTTGCATATGTATATCCATGGCCATTGTCCACCATGGCTATGTGCTGACAA 1081
Db |||||||
QY 1082 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTTTTCAATGCAATGAGTGGCAG 1141
Db |||||||
QY 876 ATGTGGCCTACTTTACGACCATTAATGCTGAGGAGCTGCTTTTCAATGCAATGAGTGGCAG 935
Db |||||||
QY 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCGATCTTTGTTGCC 1201
Db |||||||

Db 936 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCGATCTTTGTTGCC- 994
QY 1202 TCTCTGCTTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATCTATGTTG 1261
Db 995 ----- 994
QY 1262 CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCAATGTCGCGAAGCACACTC 1321
Db 995 -----CCCTC 999
QY 1322 CTCTACCAGCTGTATTGTTTTCACCCCTTTTGACAATGATAATGCTCTTCTCTGGAGACC 1381
Db |||||||
QY 1000 CTCTACCAGCTGTATTGTTTTCACCCCTTTTGACAATGATAATGCTCTTCTCTGGAGACC 1059
QY 1382 TCGACAGTCTTTTGAATTTCTCAGTTTTCAGGTTGGCTTTTATTGGGCTGGCAGTTG 1441
Db |||||||
QY 1060 TCGACAGTCTTTTGAATTTCTCAGTTTTCAGGTTGGCTTTTATTGGGCTGGCAGTTG 1119
QY 1442 CTGGGCTGATTATCTTCGATACAAATGCCAGATATGATCGTCTTCTTCAAGGTGCCAC 1501
Db |||||||
QY 1120 CTGGGCTGATTATCTTCGATACAAATGCCAGATATGATCGTCTTCTTCAAGGTGCCAC 1179
QY 1502 TGTTCATCCAGCTTTTGTTCCTTCACATGCCTTTCATGTTGCCCTTTCCTCTATT 1561
Db |||||||
QY 1180 TGTTCATCCAGCTTTTGTTCCTTCACATGCCTTTCATGTTGCCCTTTCCTCTATT 1239
QY 1562 CGGACCCATTAGTACAGGGATTGGCTTC 1590
Db |||||||
QY 1240 CGGACCCATTAGTACAGGGATTGGCTTC 1268
Db |||||||

RESULT 12

US-10-163-866-32
; Sequence 32, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-32

Query Match 20.1%; Score 449; DB 9; Length 520;
Best Local Similarity 90.5%; Pred. No. 1.2e-110;
Matches 513; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 182 GAAAGCCTGTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGGAATGTTAACGGGA 241
Db |||||||
QY 242 GGCTGCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTTCAGGAAGAGAGC 301
Db |||||||
QY 68 GGCTGCCTTCCCTGGGCAACAGGAGCCACCTGGGCGAG----- 105
QY 302 CCTTTTCAGGAAGAGAGACGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA 361
Db |||||||

Query Match	61.3%;	Score 1373.4;	DB 9;	Length 2000;				
Best Local Similarity	96.0%;	Pred. No. 0;						
Matches 1441;	Conservative	0;	Mismatches	6;	Indels	54;	Gaps	1;
QY	182	GAAGCGCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGGGAATGTTAAACGGGA	241					
Db	143	GAAGCGCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGGGAATGTTAAACGGGA	202					
QY	242	GGCTGCCCTTCCCTGGGCAACAAGGAGCCACCTGGGAGGAGCGCCTTTTCAGGAAGAGACG	301					
Db	203	GGCTGCCCTTCCCTGGGCAACAAGGAGCCACCTGGGAGGAGCGCCTTTTCAGGAAGAGACG	240					
QY	302	CCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCA	361					
Db	241	-----GAGAAAGTGCAGCTGAAGAGGAAAGTCA	268					
QY	362	CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT	421					
Db	269	CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT	328					
QY	422	CTCCTAAGGGGTGCTCCAGAACACGGGCAGCGTGGGCATGCTCTTGACCATCTGGACGG	481					
Db	329	CTCCTAAGGGGTGCTCCAGAACACGGGCAGCGTGGGCATGCTCTTGACCATCTGGACGG	388					
QY	482	TGTGTGGGGTCCCTGCACACTATTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAA	541					
Db	389	TGTGTGGGGTCCCTGCACACTATTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAA	448					
QY	542	AGAAATCTGGAGGTCAATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTG	601					
Db	449	AGAAATCTGGAGGTCAATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCTTTTG	508					
QY	602	TACGAGTCTGGGTGGAACCTCCTCATAATACGGCCTGCAGCTACTGCTGTGATATCCCTGG	661					
Db	509	TACGAGTCTGGGTGGAACCTCCTCATAATACGGCCTGCAGCTACTGCTGTGATATCCCTGG	568					
QY	662	CATTGGACGCTACATTCTTGGAAACCATTTTTTATCAATGTGAAATCCCTGAACTTGCGA	721					
Db	569	CATTGGACGCTACATTCTTGGAAACCATTTTTTATCAATGTGAAATCCCTGAACTTGCGA	628					
QY	722	TCAAGCTCATACAGCTGTGGGCATAACTGTAGTGTGCTCTTAAATAGCATGAGTGTCA	781					
Db	629	TCAAGCTCATACAGCTGTGGGCATAACTGTAGTGTGCTCTTAAATAGCATGAGTGTCA	688					
QY	782	GCTGGAGGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA	841					
Db	689	GCTGGAGGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA	748					
QY	842	TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGAGAACTTTAAAGACGCCT	901					
Db	749	TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGAGAACTTTAAAGACGCCT	808					
QY	902	TTTCAGGAAGAGATTCAGTATTACGGCGTTGCCACTGGCTTTTATTATGGAATGTATG	961					
Db	809	TTTCAGGAAGAGATTCAGTATTACGGCGTTGCCACTGGCTTTTATTATGGAATGTATG	868					
QY	962	CATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA	1021					
Db	869	CATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA	928					

```

RESULT 11
US-10-163-866-53
; Sequence 53, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 1268

```

;; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
;; FILE REFERENCE: EX02-080C
;; CURRENT APPLICATION NUMBER: US/10/163,866
;; PRIOR FILING DATE: 2002-06-05
;; PRIOR APPLICATION NUMBER: US 60/296,076
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: US 60/328,605
;; PRIOR FILING DATE: 2001-10-10
;; PRIOR APPLICATION NUMBER: US 60/338,733
;; PRIOR FILING DATE: 2001-10-22
;; PRIOR APPLICATION NUMBER: US 60/357,253
;; PRIOR FILING DATE: 2002-02-15
;; PRIOR APPLICATION NUMBER: US 60/357,600
;; PRIOR FILING DATE: 2002-02-15
;; NUMBER OF SEQ ID NOS: 54
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 52
;; LENGTH: 1528
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-163-866-52

Query Match 63.9%; Score 1431.8; DB 9; Length 1528;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1497; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY	182	GAAGGCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGAGGGGAAATGTTAACGGGA	241
Db	30	GAAGGCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGAGGGGAAATGTTAACGGGA	89
QY	242	GGCTGCCTTCCCTGGGCAACAAGGAGGACCTGGGAGGAGGCGCTTTTCAGGAAGAGACG	301
Db	90	GGCTGCCTTCCCTGGGCAACAAGGAGGACCTGGGAGGAGGCGCTTTTCAGGAAGAGACG	127
QY	302	CCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCA	361
Db	128	-----GAGAAAGTGCAGCTGAAGAGAAAGTCA	155
QY	362	CTTTACTGAGGGGAGTCTCCATATATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT	421
Db	156	CTTTACTGAGGGGAGTCTCCATATATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT	215
QY	422	CTCCTAAGGGCGTCTCCAGAACACGGGAGCGTGGGCATGTCTCTGACCATCTGGACGG	481
Db	216	CTCCTAAGGGCGTCTCCAGAACACGGGAGCGTGGGCATGTCTCTGACCATCTGGACGG	275
QY	482	TGTTGGGGTCTGTGCTACATATTTGGAGCTTTTGCTATGCTGAATTTGGAAACAATAAA	541
Db	276	TGTTGGGGTCTGTGCTACATATTTGGAGCTTTTGCTATGCTGAATTTGGAAACAATAAA	335
QY	542	AGAAATCTGGAGGTCATTACACATATATTTGGAGCTTTTGCTATGCTGAATTTGGAAACAATAAA	601
Db	336	AGAAATCTGGAGGTCATTACACATATATTTGGAGCTTTTGCTATGCTGAATTTGGAAACAATAAA	395
QY	602	TACGAGTCTGGGTGGAACTCCTCATAATACGCGCTGCAGCTACTGCTGTGATATCCCTGG	661
Db	396	TACGAGTCTGGGTGGAACTCCTCATAATACGCGCTGCAGCTACTGCTGTGATATCCCTGG	455
QY	662	CATTGGACGCTACATCTGGAACCAATTTTATTCAATGTGAAATCCCTGAACCTTGGCA	721
Db	456	CATTGGACGCTACATCTGGAACCAATTTTATTCAATGTGAAATCCCTGAACCTTGGCA	515
QY	722	TCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGATGGTCTTAAATAGCATGAGTGCA	781
Db	516	TCAAGCTCATTACAGCTGTGGGCATAACTGTAGTGATGGTCTTAAATAGCATGAGTGCA	575
QY	782	GCTGGAGCGCGGATCCAGATTTCTTAACCTTTTGCAGCTCACAGCAATTTCTGATAA	841
Db	576	GCTGGAGCGCGGATCCAGATTTCTTAACCTTTTGCAGCTCACAGCAATTTCTGATAA	635
QY	842	TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACAGCAACTTTAAAGACGCGCT	901
Db	636	TTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCACAGCAACTTTAAAGACGCGCT	695

QY	902	TTTCAGGAAGAGATTCAACTATTACGCGGTTGCCACTGGCTTTTATATATGAATGTATG	961
Db	696	TTTCAGGAAGAGATTCAAGTATTACGCGGTTGCCACTGGCTTTTATATATGAATGTATG	755
QY	962	CATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA	1021
Db	756	CATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA	815
QY	1022	CCATTCCCTTTCGCAATATGTATATCCATGGCCATTGTACCATTTGGCTATGTGCTGACAA	1081
Db	816	CCATTCCCTTTCGCAATATGTATATCCATGGCCATTGTACCATTTGGCTATGTGCTGACAA	875
QY	1082	ATGTGGCCTACTTACGACCATTAATGTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAG	1141
Db	876	ATGTGGCCTACTTACGACCATTAATGTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAG	935
QY	1142	TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC	1201
Db	936	TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC	995
QY	1202	TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATCTATGTTG	1261
Db	996	TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATCTATGTTG	1055
QY	1262	CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACACTC	1321
Db	1056	CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACACTC	1115
QY	1322	CTCTACAGCTGTTATTTGTTTGCACCTTTGCAATGATAATGCTCTCTCTGGAGACC	1381
Db	1116	CTCTACAGCTGTTATTTGTTTGCACCTTTGCAATGATAATGCTCTCTCTGGAGACC	1175
QY	1382	TCGACAGCTTTTGAATTTCTCAGTTTTCGAGGTGGCTTTTATTTGGGCTGGCAGTTG	1441
Db	1176	TCGACAGCTTTTGAATTTCTCAGTTTTCGAGGTGGCTTTTATTTGGGCTGGCAGTTG	1235
QY	1442	CTGGGCTGATTTATCTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGGCAC	1501
Db	1236	CTGGGCTGATTTATCTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGGCAC	1295
QY	1502	TGTTTATCCAGCTTTGTTTCCCTTACATGCTCTTTCATGCTGCTTCCCTCTATT	1561
Db	1296	TGTTTATCCAGCTTTGTTTCCCTTACATGCTCTTTCATGCTGCTTCCCTCTATT	1355
QY	1562	CGGACCCATTAGTACAGGATTGGCTTCGTCATCATCTGACTGGAGTCCCTGCGTATT	1621
Db	1356	CGGACCCATTAGTACAGGATTGGCTTCGTCATCATCTGACTGGAGTCCCTGCGTATT	1415
QY	1622	ATCTCTTTATATATGGGACAAAGAACCCAGGTGGTTTGAATAATGTCAGAGAAAAATAA	1681
Db	1416	ATCTCTTTATATATGGGACAAAGAACCCAGGTGGTTTGAATAATGTCAGAGAAAAATAA	1475
QY	1682	CCAGAACATTACAAATAATCTGGAAGTTGTACCAGAGAGAGATAAGTTATGA	1734
Db	1476	CCAGAACATTACAAATAATCTGGAAGTTGTACCAGAGAGAGATAAGTTATGA	1528

RESULT 10

US-10-163-866-34
; Sequence 34, Application US/10/163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733

; LENGTH: 2482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-29

Query Match 84.3%; Score 1888.2; DB 9; Length 2482;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1965; Conservative 0; Mismatches 3; Indels 55; Gaps 2;

QY	182	GAAGCCTGTTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGAAATGTTAAACGGGA	241
Db	239	GAAGCCTGTTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGAAATGTTAAACGGGA	298
QY	242	GGCTGCTCCCTGGGCAACAGGAGCCACCTGGGCAGGACGCCCTTTTCAGGAAGAGACG	301
Db	299	GGCTGCTCCCTGGGCAACAGGAGCCACCTGGGCAG-----	336
QY	302	CCCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGCAAAAGTCAGCTGAAGAGAAAGTCA	361
Db	337	-----GAGAAAGTCAGCTGAAGAGAAAGTCA	364
QY	362	CTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCACCATTTGGAGCAGGAATCTTCATCT	421
Db	365	CTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCACCATTTGGAGCAGGAATCTTCATCT	424
QY	422	CTCCTAAGGGCTGCTCCAGAACACGGGCAGCGTGGGCATGTCTCTGACCATCTGGACGG	481
Db	425	CTCCTAAGGGCTGCTCCAGAACACGGGCAGCGTGGGCATGTCTCTGACCATCTGGACGG	484
QY	482	TGTGTGGGGTCCCTGTCACCTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACTATAA	541
Db	485	TATGTGGGGTCCCTGTCACCTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACTATAA	544
QY	542	AGAAATCTGGAGGTCAATACACATATATTTGGGAAGTCTTTTGGTCCATTACCAGCTTTTG	601
Db	545	AGAAATCTGGAGGTCAATACACATATATTTGGGAAGTCTTTTGGTCCATTACCAGCTTTTG	604
QY	602	TACGAGTCTGGGTGGAACTCCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG	661
Db	605	TACGAGTCTGGGTGGAACTCCCTCATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG	664
QY	662	CATTGGACGCTACATCTGGAACTCTTATTTATCAATGTGAATCCCTGAACCTTGCGA	721
Db	665	CATTGGACGCTACATCTGGAACTCTTATTTATCAATGTGAATCCCTGAACCTTGCGA	724
QY	722	TCAAGCTCATACAGCTGTGGGCATAAAGTGTAGTGGTCCCTAAATAGCATGAGTGCA	781
Db	725	TCAAGCTCATACAGCTGTGGGCATAAAGTGTAGTGGTCCCTAAATAGCATGAGTGCA	784
QY	782	GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATCTTGATAA	841
Db	785	GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATCTTGATAA	844
QY	842	TTATAGTCCCTGGAGTTATGCAAGTAAATTAAGGTCAAACGCAGAACTTTAAAGACGCCCT	901
Db	845	TTATAGTCCCTGGAGTTATGCAAGTAAATTAAGGTCAAACGCAGAACTTTAAAGACGCCCT	904
QY	902	TTTCAGGAAGAGATCAAGTATTAACGGGTTGCCACTGGCTTTTATTTATGGAATGTATG	961
Db	905	TTTCAGGAAGAGATCAAGTATTAACGGGTTGCCACTGGCTTTTATTTATGGAATGTATG	964
QY	962	CATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAACTAGAAACCCCTGAAAAA	1021
Db	965	CATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAACTAGAAACCCCTGAAAAA	1024
QY	1022	CCATTCCTCCCTTGCATATGTATATCCATGGCCATTTGTCACCATTTGGCTATGTCTGACAA	1081
Db	1025	CCATTCCTCCCTTGCATATGTATATCCATGGCCATTTGTCACCATTTGGCTATGTCTGACAA	1084
QY	1082	ATGTGGCCTACTTTACGACCATTAATGCTGTAGAGAGCTGCTGCTTTCAAAATGCAGTGGCAG	1141
Db	1085	ATGTGGCCTACTTTACGACCATTAATGCTGTAGAGAGCTGCTGCTTTCAAAATGCAGTGGCAG	1144

QY	1142	TGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTCCGATCTTTGTTGCC	1201
Db	1145	TGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTCCGATCTTTGTTGCC	1204
QY	1202	TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTCTCCAGGTTATCTATGTTG	1261
Db	1205	TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTCTCCAGGTTATCTATGTTG	1264
QY	1262	CGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTC	1321
Db	1265	CGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTC	1324
QY	1322	CTCTACCCAGCTGTTATGTTTTCACCCCTTTTGACAATGATAATGCTCTTCTCTGGAGACC	1381
Db	1325	CTCTACCCAGCTGTTATGTTTTCACCCCTTTTGACAATGATAATGCTCTTCTCTGGAGACC	1384
QY	1382	TCGACAGTCTTTTGAAATTTCTCCTCAGTTTTCGCCAGGTGGCTTTTATTTGGGCTGGCAGTTG	1441
Db	1385	TCGACAGTCTTTTGAAATTTCTCCTCAGTTTTCGCCAGGTGGCTTTTATTTGGGCTGGCAGTTG	1444
QY	1442	CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTTCAAGGTGCCAC	1501
Db	1445	CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTTCAAGGTGCCAC	1504
QY	1502	TGTTATCCAGCTTTTCTTTCCTTTCACATGCTCTTTCATGTTGCTCTTCCCTCTATT	1561
Db	1505	TGTTATCCAGCTTTTCTTTCCTTTCACATGCTCTTTCATGTTGCTCTTCCCTCTATT	1564
QY	1562	CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCCTCTGACTGGAGTCCCTGCGTATT	1621
Db	1565	CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCCTCTGACTGGAGTCCCTGCGTATT	1624
QY	1622	ATCTCTTTATATATGGGACAAAGAACCCAGGTGTTTAGAATAATGTACAGAGAAATAA	1681
Db	1625	ATCTCTTTATATATGGGACAAAGAACCCAGGTGTTTAGAATAATGTACAGAGAAATAA	1684
QY	1682	CCAGAACATTAACAATAATCTGGAAGTTGTACCAGAGAAAGATAAGTTATGAACATAATG	1741
Db	1685	CCAGAACATTAACAATAATCTGGAAGTTGTACCAGAGAAAGATAAGTTATGAACATAATG	1744
QY	1742	GACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAATAAGGATTTTACTTCATT	1801
Db	1745	GACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAATAAGGATTTTACTTCATT	1804
QY	1802	TTCTGAAAGCTTAGAGAAATTACAACCTTTGGTGATAAACAAGAGAGTCAAGTTATTTTAT	1861
Db	1805	TTCTGAAAGCTTAGAGAAATTACAACCTTTGGTGATAAACAAGAGAGTCAAGTTATTTTAT	1864
QY	1862	TCATATATTTTAGCATATTCGAACCTAATTTCTAAGAAATTTAGTTATAAATCTATGTAAT	1921
Db	1865	TCATATATTTTAGCATATTCGAACCTAATTTCTAAGAAATTTAGTTATAAATCTATGTAAT	1924
QY	1922	TATAGAAAGTGAATATGAGTTATTTCTATGAGTGCACAAATTTCTGAGTCTCTGATACCT	1981
Db	1925	TATAGAAAGTGAATATGAGTTATTTCTATGAGTGCACAAATTTCTGAGTCTCTGATACCT	1984
QY	1982	ACCTATTTGGGTTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCAATTTCTTACAACATA	2041
Db	1985	ACCTATTTGGGTTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCAATTTCTTACAACATA	2044
QY	2042	TGTTAGCACGGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTATATATGGGTTT	2101
Db	2045	TGTTAGCACGGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTATATATGGGTTT	2104
QY	2102	TGTAAGAGATGTTTACACACTACAGATGTCTATCTGTAAGAAAGTGTTCATCTCTGA	2161
Db	2105	TGT-AAGATGTTTACACACTACAGATGTCTATCTGTAAGAAAGTGTTCATCTCTGA	2163
QY	2162	AAAAAGCATACATCATGATTTATGGCAAGAGGAGAGAGGTA	2204
Db	2164	AAAAAGCATACATCATGATTTATGGCAAGAGGAGAGAGAGAA	2206

Db 134 CTTTACTGAGGGGAGTCTCCATATATCATTTGGCACCATCATTTGGAGCAGGAATCTTTCATCT 193

Qy 422 CTCCTAAGGGCGTGTCCAGAACACGGGCGAGCGTGGGCATGTCTGTGACCATCTGGACGG 481

Db 194 CTCCTAAGGGCGTGTCCAGAACACGGGCGAGCGTGGGCATGTCTGTGACCATCTGGACGG 253

Qy 482 TGTGTGGGTCCTGTCTACATATTTGGAGCTTTGTCTTATGCTGAATGGGAACAACATATAA 541

Db 254 TGTGTGGGTCCTGTCTACATATTTGGAGCTTTGTCTTATGCTGAATGGGAACAACATATAA 313

Qy 542 AGAAATCTGGAGGTCATTACACATATATTTGGAAGCTTTGGTCCATTACCCAGCTTTTG 601

Db 314 AGAAATCTGGAGGTCATTACACATATATTTGGAAGCTTTGGTCCATTACCCAGCTTTTG 373

Qy 602 TACGAGTCTGGGTGGAATCCTCATATACGCCCTGCAGTACTGCTGTGATATCCCTGG 661

Db 374 TACGAGTCTGGGTGGAATCCTCATATACGCCCTGCAGTACTGCTGTGATATCCCTGG 433

Qy 662 CATTTGGACGCTACATCTCGAACCATTTTTATCAATGTGAATCCCTGAACCTTGCGA 721

Db 434 CATTTGGACGCTACATCTCGAACCATTTTTATCAATGTGAATCCCTGAACCTTGCGA 493

Qy 722 TCAAGCTCATACAGCTGTGGGCATAACTGTAGTATGCTTAAAGCTCAGCAATCTGATAA 781

Db 494 TCAAGCTCATACAGCTGTGGGCATAACTGTAGTATGCTTAAAGCTCAGCAATCTGATAA 553

Qy 782 GCTGGAGCGCCGGATCCAGATTTCTTAACTTTTGAAGCTCAGCAATCTGATAA 841

Db 554 GCTGGAGCGCCGGATCCAGATTTCTTAACTTTTGAAGCTCAGCAATCTGATAA 613

Qy 842 TTATAGTCCCTGGAGTATGCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGACGCGT 901

Db 614 TTATAGTCCCTGGAGTATGCAGCTAATTAAGGTCAAACGCAGAACTTTAAAGACGCGT 673

Qy 902 TTTTCAAGGAGAGATTTCAAGTATACGCGGTGCCACTGGCTTTTATTTATGGAATGATG 961

Db 674 TTTTCAAGGAGAGATTTCAAGTATACGCGGTGCCACTGGCTTTTATTTATGGAATGATG 733

Qy 962 CATATGCTGGCTGGTCTTACCTCAACTTTGTACTGAAGAGTGAAGAACCCCTGAAAAA 1021

Db 734 CATATGCTGGCTGGTCTTACCTCAACTTTGTACTGAAGAGTGAAGAACCCCTGAAAAA 793

Qy 1022 CCATTCCCTTGCATATATATCCATGGCCATTTCTCATAGCAGTTCCGATCTTTGTTGCC 1081

Db 794 CCATTCCCTTGCATATATATATCCATGGCCATTTCTCATAGCAGTTCCGATCTTTGTTGCC 853

Qy 1082 ATGTGGCCCTACTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAATGCAAGTGGCAG 1141

Db 854 ATGTGGCCCTACTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAATGCAAGTGGCAG 913

Qy 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCCGATCTTTGTTGCC 1201

Db 914 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCCGATCTTTGTTGCC 973

Qy 1202 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGTTATCTATGTTG 1261

Db 974 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGTTATCTATGTTG 1033

Qy 1262 CGTCTGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACATC 1321

Db 1034 CGTCTGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACATC 1093

Qy 1322 CTCTACAGCTGTTATTTTGCACCCTTTGACAAATGATAATGCTCTCTCTGGAGACC 1381

Db 1094 CTCTACAGCTGTTATTTTGCACCCTTTGACAAATGATAATGCTCTCTCTGGAGACC 1153

Qy 1382 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGCTGGCAGTTG 1441

Db 1154 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGCTGGCAGTTG 1213

Qy 1442 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTTTCAAGTGCCAC 1501

Db 1214 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTTTCAAGTGCCAC 1273

Qy 1502 TGTTCATCCAGCTTTGTTTTCCCTTACATGCCTCTTCATGGTTGCCCTTCCCTCTATT 1561

Db 1274 TGTTCATCCAGCTTTGTTTTCCCTTACATGCCTCTTCATGGTTGCCCTTCCCTCTATT 1333

Qy 1562 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATT 1621

Db 1334 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATT 1393

Qy 1622 ATCTCTTTTATATATGGGACAAAGAAACCCAGGTGTTTGAATAATGTAGAGATAA 1681

Db 1394 ATCTCTTTTATATATGGGACAAAGAAACCCAGGTGTTTGAATAATGTAGAGATAA 1453

Qy 1682 CCAGAACATTTACAATAATACTGGAAGTTGTACCAGAGAGATAAGTATGAACATAA 1741

Db 1454 CCAGAACATTTACAATAATACTGGAAGTTGTACCAGAGAGATAAGTATGAACATAA 1513

Qy 1742 GACTTGAGATCTTGGCAATCTGCCAAGTGGGAGACACAAATAAGGATTTTACTTCATT 1801

Db 1514 GACTTGAGATCTTGGCAATCTGCCAAGTGGGAGACACAAATAAGGATTTTACTTCATT 1573

Qy 1802 TTCTGAAAGCTAGAGAAATTAACAATTTGGTGATATAAACAAGAGAGTATTTTAT 1861

Db 1574 TTCTGAAAGCTAGAGAAATTAACAATTTGGTGATATAAACAAGAGAGTATTTTAT 1633

Qy 1862 TCATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATAACTCTATGATG 1921

Db 1634 TCATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATAACTCTATGATG 1693

Qy 1922 TATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTGAGTCTCTGATACCT 1981

Db 1694 TATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCACAAATTTGAGTCTCTGATACCT 1753

Qy 1982 ACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGTTCAATTTCTACAACATA 2041

Db 1754 ACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGTTCAATTTCTACAACATA 1813

Qy 2042 TGTAGCACGGCAAGAACTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 2101

Db 1814 TGTAGCACGGCAAGAACTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 1873

Qy 2102 TGTAAAGATGTTTACACACTACAGATGTTCTACTATGTTGTTCAATTTCTCAATTTCTGA 2161

Db 1874 TGTAAAGATGTTTACACACTACAGATGTTCTACTATGTTGTTCAATTTCTCAATTTCTGA 1933

Qy 2162 AAAAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGGTA 2204

Db 1934 AAAAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGAGAA 1976

RESULT 6

US-10-163-866-29

; Sequence 29, Application US/10163866

; Publication No. US20030027188A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/10/163,866

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: US 60/357,600

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 29

Db 1747 TCTGAAAGTCTAGAGAAATTACAACTTTGGTGATATAAAAGGAGTCAGTTATTTATTT 1806
QY 1863 CATATATTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATATACTCTATGTAGTT 1922
Db 1807 CATATATTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATATACTCTATGTAGTT 1866
QY 1923 ATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCAAAATTTCTTGTGATCTCTGATACCTA 1982
Db 1867 ATAGAAAGTGAATATGCAGTTATTTCTATGAGTCGCAAAATTTCTTGTGATCTCTGATACCTA 1926
QY 1983 CCTATTGGGGTTAGGAGAAAGACTAGACAATTTACTATGTGGTCATTTCTCTACAACATAT 2042
Db 1927 CCTATTGGGGTTAGGAGAAAGACTAGACAATTTACTATGTGGTCATTTCTCTACAACATAT 1986
QY 2043 GTTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 2102
Db 1987 GTTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 2046
QY 2103 GTAAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAAGTGTTTTCAATTCGTGAA 2162
Db 2047 GTAAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAAGTGTTTTCAATTCGTGAA 2106
QY 2163 AAAAGCATACATCATGATTATGGCAAGAGGAGAGAGGTA 2204
Db 2107 AAAAGCATACATCATGATTATGGCAAGAGGAGAGAGAGAA 2148

RESULT 5

US-10-163-866-31
; Sequence 31, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXPLIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 3144
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-31

Query Match 84.9%; Score 1900.2; DB 9; Length 3144;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1966; Conservative 0; Mismatches 3; Indels 54; Gaps 1;

QY 182 GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGGA 241
Db 8 GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGGA 67
QY 242 GGCTGCTTCCCTGGGCAACAGGACCCACCTGGGCGAGGACGCTTTTCAGGAAGAGACG 301
Db 68 GGCTGCTTCCCTGGGCAACAGGACCCACCTGGGCGAG----- 105
QY 302 CCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAGAAAGTCA 361
Db 106 -----GAGAAAGTGCAGCTGAAGAGAGAAAGTCA 133
QY 362 CTTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATTGAGCAGGAAATCTTCATCT 421

Db 667 CAAGCTCATACAGCTGTGGCATAACTAGTGTGTCCTAAATAGCATGAGTGTGAG 726
QY 783 CTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAAT 842
Db 727 CTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAAT 786
QY 843 TATAGTCCCTGGAGTTATGCAGCTAATTAAGAGTCAACCGCAGAACTTTAAAGACGCCTT 902
Db 787 TATAGTCCCTGGAGTTATGCAGCTAATTAAGAGTCAACCGCAGAACTTTAAAGACGCCTT 846
QY 903 TTCAGGAAGAGATTCAAGATTATACGCGGTTGCCACTGGCTTTTATATGGAATGTATGC 962
Db 847 TTCAGGAAGAGATTCAAGATTATACGCGGTTGCCACTGGCTTTTATATGGAATGTATGC 906
QY 963 ATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAGTAGAAAACCCCTGAAAAAAC 1022
Db 907 ATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAGTAGAAAACCCCTGAAAAAAC 966
QY 1023 CATTCCTTGCATATGTATATCCATGGCCATTGTCCACCATTTGGCTATGTGCTGACAAA 1082
Db 967 CATTCCTTGCATATGTATATCCATGGCCATTGTCCACCATTTGGCTATGTGCTGACAAA 1026
QY 1083 TGTGGCCTACTTTACGACCATTTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGT 1142
Db 1027 TGTGGCCTACTTTACGACCATTTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGT 1086
QY 1143 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCGGATCTTTTGTGCCCT 1202
Db 1087 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCGGATCTTTTGTGCCCT 1146
QY 1203 CTCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTGC 1262
Db 1147 CTCCTGCTTTGGCTCCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTGC 1206
QY 1263 GTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCTCCGCAAGCACACTCC 1322
Db 1207 GTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCTCCGCAAGCACACTCC 1266
QY 1323 TCTACACAGCTGTTATTGTTTGCACCTTTGACAAATGATAATGCTTCTCTGGAGACCT 1382
Db 1267 TCTACACAGCTGTTATTGTTTGCACCTTTGACAAATGATAATGCTTCTCTGGAGACCT 1326
QY 1383 CGACAGTCTTTTGAATTTCTCAGTGTGTTGCCAGGTGGCTTTTATTTGGGCTGGCAGTTGC 1442
Db 1327 CGACAGTCTTTTGAATTTCTCAGTGTGTTGCCAGGTGGCTTTTATTTGGGCTGGCAGTTGC 1386
QY 1443 TGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCACT 1502
Db 1387 TGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCACT 1446
QY 1503 GTTCATCCAGCTTTGTTTTCCTTCACATGCCTTCATGTTGGCTTCCCTCTATTTC 1562
Db 1447 GTTCATCCAGCTTTGTTTTCCTTCACATGCCTTCATGTTGGCTTCCCTCTATTTC 1506
QY 1563 GGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTA 1622
Db 1507 GGACCCATTTAGTACAGGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCGTATTA 1566
QY 1623 TCTCTTTATTTATGGGACAAAGAACCCAGGTGGTTTGAATAATGTGACAGAGAAATAAC 1682
Db 1567 TCTCTTTATTTATGGGACAAAGAACCCAGGTGGTTTGAATAATGTGACAGAGAAATAAC 1626
QY 1683 CAGAACATTACAAATAACTGGAAGTTGTACCAGAGAGATAAGTTATGAACATAATGG 1742
Db 1627 CAGAACATTACAAATAACTGGAAGTTGTACCAGAGAGATAAGTTATGAACATAATGG 1686
QY 1743 ACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGATTTTACTTCAATT 1802
Db 1687 ACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGATTTTACTTCAATT 1746
QY 1803 TCTGAAAGTCTAGAGAAATTACAACTTTGGTGATAAAACAAAAGGAGTCAATTTATTTAT 1862

; APPLICANT: Fling, Steven P.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Secret, Heather
 ; APPLICANT: Indirias, Carol Yoseph
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Elliot, Mark
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Kalos, Michael D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; THE THERAPY AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.475C9
 ; CURRENT APPLICATION NUMBER: US/09/738,973
 ; CURRENT FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 587
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 440
 ; LENGTH: 2239
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-738-973-440

Query Match 100.0%; Score 2239; DB 10; Length 2239;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAGGTTGAAGTACAGAGATCATGCCAGCCTGGGTGACAGTGGAGTCTGTCTCAAAAC	60
Db	1	GGAGGTTGAAGTACAGAGATCATGCCAGCCTGGGTGACAGTGGAGTCTGTCTCAAAAC	60
QY	61	AGAAATTAAGGAAAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCGCAATTGTGG	120
Db	61	AGAAATTAAGGAAAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCGCAATTGTGG	120
QY	121	CATAGATTTATCATATTTCTGGATTTTGGATTTCTTTGTTTCTCATCTGATTTCA	180
Db	121	CATAGATTTATCATATTTCTGGATTTTGGATTTCTTTGTTTCTCATCTGATTTCA	180
QY	181	GGAAAGCCTGTGTGTCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGG	240
Db	181	GGAAAGCCTGTGTGTCCACCATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGG	240
QY	241	AGGCTGCCTTCCCTGGGCAACAAAGGAGCCACCTGGGCGAGGAGCCCTTTTCAGGAAGAGAC	300
Db	241	AGGCTGCCTTCCCTGGGCAACAAAGGAGCCACCTGGGCGAGGAGCCCTTTTCAGGAAGAGAC	300
QY	301	GCCTTTTCAGGAAGAGAGCGCCTTTTCAGGAAGAGAGAAAGTGCAGTGAAGAGGAAAGTC	360
Db	301	GCCTTTTCAGGAAGAGAGCGCCTTTTCAGGAAGAGAGAAAGTGCAGTGAAGAGGAAAGTC	360
QY	361	ACTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATC	420
Db	361	ACTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATC	420
QY	421	TCTCCTAAGGGCGTCTCCAGAACACGGGCGAGCGTGGGCGATGTCTTGACCATCTGGACG	480
Db	421	TCTCCTAAGGGCGTCTCCAGAACACGGGCGAGCGTGGGCGATGTCTTGACCATCTGGACG	480
QY	481	GTGTGTGGGGTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAATA	540
Db	481	GTGTGTGGGGTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAATA	540
QY	541	AAGAAATCTGGAGTCAATTACACATATATTTTGGAGTCTTTGGTCCATTACAGCTTTT	600
Db	541	AAGAAATCTGGAGTCAATTACACATATATTTTGGAGTCTTTGGTCCATTACAGCTTTT	600
QY	601	GTACGAGTCTGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTG	660
Db	601	GTACGAGTCTGGGTGGAACCTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTG	660
QY	661	GCATTTGGACGCTACATTTCTGGAACCAATTTTATCAATGTGAAATCCCTGAACTTGG	720
Db	661	GCATTTGGACGCTACATTTCTGGAACCAATTTTATCAATGTGAAATCCCTGAACTTGG	720

QY	721	ATCAAGCTCATACAGCTGTGGGCATAACTGTAGTGTCTCTAAATAGCATGAGTGT	780
Db	721	ATCAAGCTCATACAGCTGTGGGCATAACTGTAGTGTCTCTAAATAGCATGAGTGT	780
QY	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA	840
Db	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA	840
QY	841	ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGAGCAACTTTAAAGACGCC	900
Db	841	ATTATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGAGCAACTTTAAAGACGCC	900
QY	901	TTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTAATGAATGTAT	960
Db	901	TTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTAATGAATGTAT	960
QY	961	GCATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAAAGTAGAAACCCCTGAAAA	1020
Db	961	GCATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAAAGTAGAAACCCCTGAAAA	1020
QY	1021	ACCATTCCTTGCATATATATCCATGGCCATTTCCACCATTTGGCTATGTGCTGACA	1080
Db	1021	ACCATTCCTTGCATATATATCCATGGCCATTTCCACCATTTGGCTATGTGCTGACA	1080
QY	1081	AATGTGGCTACTTTACGACCATTAATGCTGAGGAGTGTCTCTCAATGCAGTGGCA	1140
Db	1081	AATGTGGCTACTTTACGACCATTAATGCTGAGGAGTGTCTCTCAATGCAGTGGCA	1140
QY	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTCCGATCTTTGTTGCC	1200
Db	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTCCGATCTTTGTTGCC	1200
QY	1201	CTCTCCTGCTTTGGCTCCATGAACGGTGTGTGTTTGTCTCTCCAGGTTATTTCTATGTT	1260
Db	1201	CTCTCCTGCTTTGGCTCCATGAACGGTGTGTGTTTGTCTCTCCAGGTTATTTCTATGTT	1260
QY	1261	GCGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGAAGCACACT	1320
Db	1261	GCGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCGAAGCACACT	1320
QY	1321	CCTCTACAGCTGTATTGTTTGCACCCCTTTTGACAATGATAATGCTCTCTCTGGAGAC	1380
Db	1321	CCTCTACAGCTGTATTGTTTGCACCCCTTTTGACAATGATAATGCTCTCTCTGGAGAC	1380
QY	1381	CTCAGCAGTCTTTTGAATTTCTCAGTTTGGCCAGGTGGCTTTTATTTGGGCTGGCAGTT	1440
Db	1381	CTCAGCAGTCTTTTGAATTTCTCAGTTTGGCCAGGTGGCTTTTATTTGGGCTGGCAGTT	1440
QY	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTCAAGGTGCCA	1500
Db	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTCAAGGTGCCA	1500
QY	1501	CTGTTTCATCCAGCTTTGTTTCTCTTCCATGCTTCCATGCTTCCCTCTAT	1560
Db	1501	CTGTTTCATCCAGCTTTGTTTCTCTTCCATGCTTCCATGCTTCCCTCTAT	1560
QY	1561	TCGACCCCATTTAGTACAGGATTTGGCTTCGTCATCATCTGACTGGAGTCCCTGCGTAT	1620
Db	1561	TCGACCCCATTTAGTACAGGATTTGGCTTCGTCATCATCTGACTGGAGTCCCTGCGTAT	1620
QY	1621	TATCTCTTTTATATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTACAGAAAAATA	1680
Db	1621	TATCTCTTTTATATATGGGACAAAGAACCCAGGTGGTTTAGAATAATGTACAGAAAAATA	1680
QY	1681	ACCAGAACATTACAAATAATCTGGAAGTGTACCAGAAAGATAAGTTATGAACATAAT	1740
Db	1681	ACCAGAACATTACAAATAATCTGGAAGTGTACCAGAAAGATAAGTTATGAACATAAT	1740
QY	1741	GGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAT	1800
Db	1741	GGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAT	1800

QY 1321 CCTCTACAGCTGTTATGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGAC 1380
Db 1321 CCTCTACAGCTGTTATGTTTGCACCCCTTTGACAAATGATAATGCTCTTCTCTGGAGAC 1380
QY 1381 CTCGACAGCTCTTTGAATTTCCCTCAGTTTTCAGGCTGGCTTTTATTGGGCTGGCAGTT 1440
Db 1381 CTCGACAGCTCTTTGAATTTCCCTCAGTTTTCAGGCTGGCTTTTATTGGGCTGGCAGTT 1440
QY 1441 GCTGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCA 1500
Db 1441 GCTGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCA 1500
QY 1501 CTGTTTCATCCAGCTTTGTTTCCCTTCACATGCCCTCTCATGGTTGCCCTTCCCTCTAT 1560
Db 1501 CTGTTTCATCCAGCTTTGTTTCCCTTCACATGCCCTCTCATGGTTGCCCTTCCCTCTAT 1560
QY 1561 TCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACCTCTGACTGGAGTCCCTGCCGTAT 1620
Db 1561 TCGGACCCATTTAGTACAGGGATTGGCTTCGTCATCACCTCTGACTGGAGTCCCTGCCGTAT 1620
QY 1621 TATCTCTTTATTATGGGACAAAGAACCCAGGTTGTACCAGAAAGATAAGTTATGAAGATA 1680
Db 1621 TATCTCTTTATTATGGGACAAAGAACCCAGGTTGTACCAGAAAGATAAGTTATGAAGATA 1680
QY 1681 ACCAGAACATTTACAAATAACTGGAAGTTGTACCAGAAAGATAAGTTATGAAGATA 1740
Db 1681 ACCAGAACATTTACAAATAACTGGAAGTTGTACCAGAAAGATAAGTTATGAAGATA 1740
QY 1741 GGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCAT 1800
Db 1741 GGACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCAT 1800
QY 1801 TTTCTGAAAGTCTAGAGAAATACAACTTTTGGTGATAAAACAAAGGAGTCACTATTTTA 1860
Db 1801 TTTCTGAAAGTCTAGAGAAATACAACTTTTGGTGATAAAACAAAGGAGTCACTATTTTA 1860
QY 1861 TTCATATATTTTAGCATATTTTGAAGTAAATTTCTAAGAAATTTAGTTATACTTATGTAG 1920
Db 1861 TTCATATATTTTAGCATATTTTGAAGTAAATTTCTAAGAAATTTAGTTATACTTATGTAG 1920
QY 1921 TTATAGAAAGTGAATATGCAGTTATCTATGAGTGGCAATTTCTTGAGTCTCTGATACC 1980
Db 1921 TTATAGAAAGTGAATATGCAGTTATCTATGAGTGGCAATTTCTTGAGTCTCTGATACC 1980
QY 1981 TACCTATTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCACTCTCTACACAT 2040
Db 1981 TACCTATTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCACTCTCTACACAT 2040
QY 2041 ATGTTAGCACGGCAAAAGACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGT 2100
Db 2041 ATGTTAGCACGGCAAAAGACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGT 2100
QY 2101 TTGTAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAAGTGTCTTCAATTCG 2160
Db 2101 TTGTAAGATGGTTTACACACTACAGATGTCTATACTGTGAAAAGTGTCTTCAATTCG 2160
QY 2161 AAAAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2161 AAAAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2221 TATTAATAAAAAAAAAAAAA 2239
Db 2221 TATTAATAAAAAAAAAAAAA 2239

RESULT 2
US-09-738-973-440
; Sequence 440, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.

QY 241 AGGCTGCCTTCCCTGGGCAACAAAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGAC 300
Db 241 AGGCTGCCTTCCCTGGGCAACAAAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGAC 300
QY 301 GCCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGTGAAGAGGAAAGTC 360
Db 301 GCCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAGAAAGTGCAGTGAAGAGGAAAGTC 360
QY 361 ACTTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATTATGGAGAGGAATCTTCATC 420
Db 361 ACTTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATTATGGAGAGGAATCTTCATC 420
QY 421 TCTCCTTAAGGGCGTGTCTCCAGAACACGGGAGCGGCGGCGATGTCTCTGACCATCTGGACG 480
Db 421 TCTCCTTAAGGGCGTGTCTCCAGAACACGGGAGCGGCGGCGATGTCTCTGACCATCTGGACG 480
QY 481 GTGTGTGGGCTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACTATA 540
Db 481 GTGTGTGGGCTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACTATA 540
QY 541 AAGAAATCTGGAGGTCAATACACATATATTTTGAAGTCTTTGGTCCATTACCAGCTTTT 600
Db 541 AAGAAATCTGGAGGTCAATACACATATATTTTGAAGTCTTTGGTCCATTACCAGCTTTT 600
QY 601 GTACGAGTCTGGTGGAACTCCTCAATAATACGCCCTGCAAGTACTGCTGTGATATCCCTG 660
Db 601 GTACGAGTCTGGTGGAACTCCTCAATAATACGCCCTGCAAGTACTGCTGTGATATCCCTG 660
QY 661 GCATTTGGACGCTACATTTCTGGAACCAATTTTATTCAAATGTAATCCCTGAACTTGGC 720
Db 661 GCATTTGGACGCTACATTTCTGGAACCAATTTTATTCAAATGTAATCCCTGAACTTGGC 720
QY 721 ATCAAGCTCATTTACAGCTGTGGGCATTAAGTGTAGTGTCTTAAATAGCATGAGTGTG 780
Db 721 ATCAAGCTCATTTACAGCTGTGGGCATTAAGTGTAGTGTCTTAAATAGCATGAGTGTG 780
QY 781 AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATA 840
Db 781 AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATA 840
QY 841 ATTAATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCAGCACTTTAAAGACGCC 900
Db 841 ATTAATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAAACGCAGCACTTTAAAGACGCC 900
QY 901 TTTTCAGGAAGAGATTTCAAGTATTAAGCGGTTGCCACTGGCTTTTATTTATGGAATGTAT 960
Db 901 TTTTCAGGAAGAGATTTCAAGTATTAAGCGGTTGCCACTGGCTTTTATTTATGGAATGTAT 960
QY 961 GCATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAAAGTAGAAAACCTGAAAAA 1020
Db 961 GCATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAAAGTAGAAAACCTGAAAAA 1020
QY 1021 ACCATTCCCTTGCATATGATATATCCATGGCCATTTGTCAACCATTTGGCTATGTGCTGACA 1080
Db 1021 ACCATTCCCTTGCATATGATATATCCATGGCCATTTGTCAACCATTTGGCTATGTGCTGACA 1080
QY 1081 AATGTGGCTTACTTTACGACCAATTAATGCTGAGGAGCTGTCTTCAAAATGCAAGTGGCA 1140
Db 1081 AATGTGGCTTACTTTACGACCAATTAATGCTGAGGAGCTGTCTTCAAAATGCAAGTGGCA 1140
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTACAGTTTCCGATCTTTGTGCC 1200
Db 1141 GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCCGATCTTTGTGCC 1200
QY 1201 CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAGGTTATCTATGTT 1260
Db 1201 CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTTGTCTGTCTCCAGGTTATCTATGTT 1260
QY 1261 GCGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCCCGCAAGCACACT 1320
Db 1261 GCGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCCCGCAAGCACACT 1320

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 06:28:33 ; Search time 232 Seconds
(without alignments) 9694.889 Million cell updates/sec

Title: US-09-667-170A-440
 Perfect score: 2239
 Sequence: 1 ggagggtgaagtgcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

Database : Published_Applications NA:

```

1: /cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			SUMMARIES		Description
	Score	Match	Length	DB	ID	
1	2239	100.0	2239	9	US-09-854-133-440	Sequence 440, Appl
2	2239	100.0	2239	10	US-09-738-973-440	Sequence 440, Appl
3	2080.8	92.9	5981	9	US-09-854-133-441	Sequence 441, Appl
4	2080.8	92.9	5981	10	US-09-738-973-441	Sequence 441, Appl
5	1900.2	84.9	3144	9	US-10-163-866-31	Sequence 31, Appl
6	1888.2	84.3	2482	9	US-10-163-866-29	Sequence 29, Appl
7	1540.8	68.8	1861	9	US-10-163-866-30	Sequence 30, Appl
8	1454.4	65.0	1542	9	US-10-163-866-33	Sequence 33, Appl
9	1431.8	63.9	1528	9	US-10-163-866-52	Sequence 52, Appl
10	1373.4	61.3	2000	9	US-10-163-866-34	Sequence 34, Appl
11	1047.4	46.8	1268	9	US-10-163-866-53	Sequence 53, Appl
12	449	20.1	520	9	US-10-163-866-32	Sequence 32, Appl
13	379.2	16.9	1581	9	US-10-163-866-10	Sequence 10, Appl
14	379.2	16.9	6295	9	US-10-163-866-8	Sequence 8, Appl
15	379.2	16.9	6296	9	US-10-163-866-9	Sequence 9, Appl
16	366.4	16.4	1582	9	US-10-163-866-12	Sequence 12, Appl
17	365.4	16.3	1609	9	US-10-163-866-5	Sequence 5, Appl
18	365.4	16.3	1621	9	US-10-163-866-7	Sequence 7, Appl
19	365.4	16.3	4559	10	US-09-919-172-61	Sequence 61, Appl

20	365.4	16.3	4559	12	US-10-044-090-555	Sequence 555, Appl
21	364.8	16.3	2098	9	US-10-163-866-11	Sequence 11, Appl
22	362.2	16.2	1541	9	US-10-163-866-51	Sequence 51, Appl
23	362.2	16.2	1559	9	US-10-163-866-3	Sequence 3, Appli
24	362.2	16.2	1593	9	US-10-163-866-1	Sequence 1, Appli
25	362.2	16.2	4539	9	US-10-163-866-2	Sequence 2, Appli
26	343.8	15.4	3728	9	US-10-163-866-14	Sequence 14, Appl
27	343.8	15.4	4237	9	US-10-163-866-15	Sequence 15, Appl
28	343.8	15.4	4237	9	US-10-163-866-20	Sequence 20, Appl
29	342.2	15.3	1657	9	US-10-163-866-16	Sequence 16, Appl
30	337.2	15.1	3717	9	US-10-163-866-13	Sequence 13, Appl
31	319.6	14.3	2529	9	US-09-815-923-15	Sequence 15, Appl
32	315.6	14.1	1595	9	US-10-163-866-27	Sequence 27, Appl
33	315.6	14.1	1918	9	US-10-163-866-26	Sequence 26, Appl
34	311.6	13.9	1962	9	US-10-163-866-28	Sequence 28, Appl
35	301	13.4	1476	9	US-10-163-866-24	Sequence 24, Appl
36	301	13.4	1654	9	US-10-163-866-22	Sequence 22, Appl
37	301	13.4	1664	9	US-10-163-866-25	Sequence 25, Appl
38	299.4	13.4	1814	9	US-10-163-866-23	Sequence 23, Appl
39	297.8	13.3	1651	9	US-10-163-866-21	Sequence 21, Appl
40	281	12.6	337	9	US-09-854-133-442	Sequence 442, App
41	281	12.6	337	10	US-09-738-973-442	Sequence 442, App
42	238.4	10.6	3984	9	US-09-954-531-151	Sequence 151, App
43	238.4	10.6	3984	9	US-10-163-866-6	Sequence 6, Appli
44	230	10.3	572	9	US-10-046-935-1307	Sequence 1307, Ap
45	230	10.3	572	9	US-09-878-178-1307	Sequence 1307, Ap

ALIGNMENTS

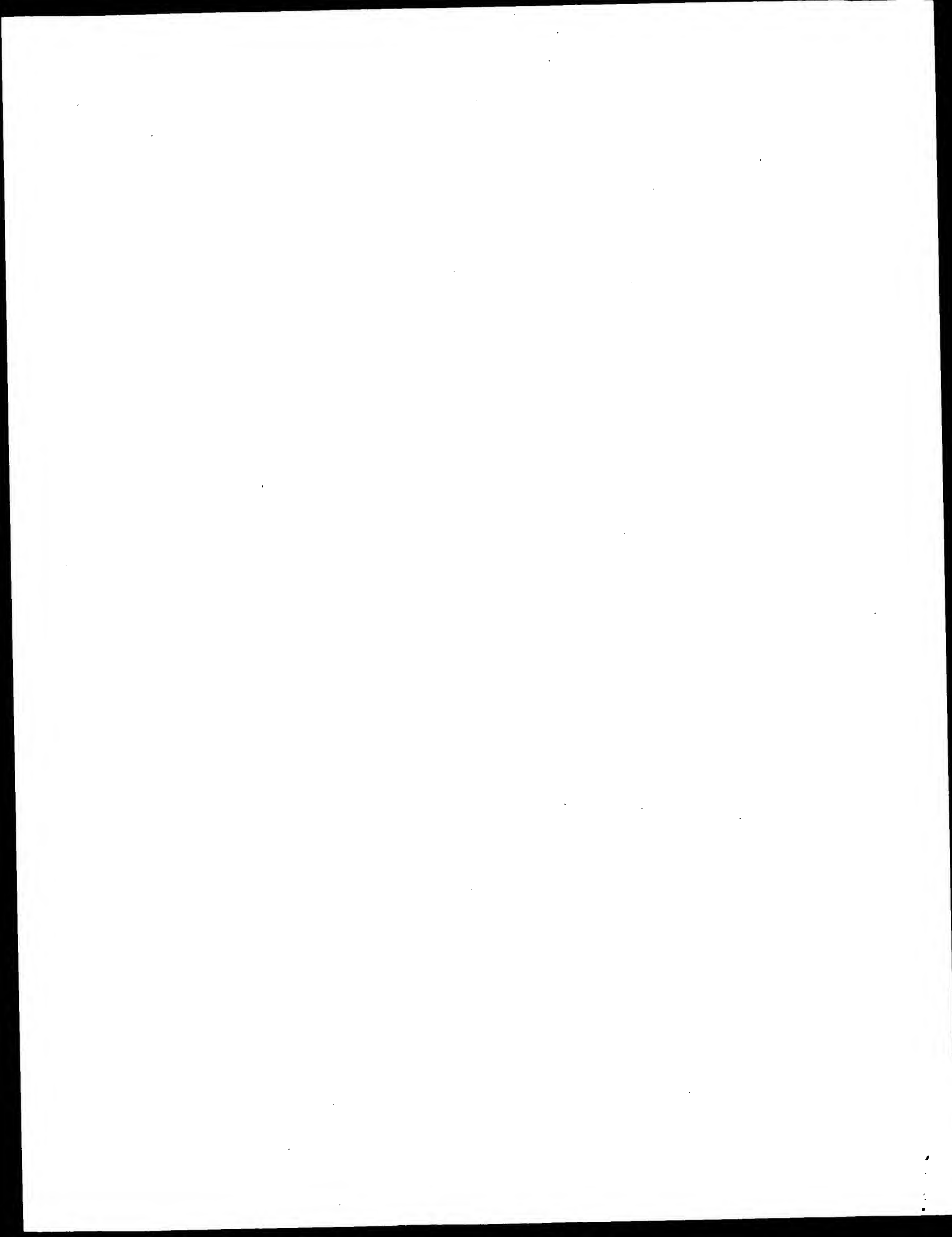
```

RESULT 1
US-09-854-133-440
; Sequence 440, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 440
; LENGTH: 2239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-440

```

Query Match	100.0%	Score 2239;	DB 9;	Length 2239;
Best Local Similarity	100.0%	Pred. NO. 0;		
Matches 2239;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGCAGACTCTGTCTCAAAC	60
Db	1	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGCAGACTCTGTCTCAAAC	60
QY	61	AGAATTAAAGCAAAAAAGAAAGAAAAGAGAGAGAGGAAATCCAGGCCAATTGTGG	120
Db	61	AGAATTAAAGCAAAAAAGAAAGAAAAGAGAGAGAGGAAATCCAGGCCAATTGTGG	120
QY	121	CATAGATTTTATCATATCTCGAATTTTITGGATTCTTTTGTTTCTCATCACTGGATTCA	180
Db	121	CATAGATTTTATCATATCTCGAATTTTITGGATTCTTTTGTTTCTCATCACTGGATTCA	180
QY	181	GAAAGCCTGTTGTGTCCACCATCTCCAAGGAGGTTAACTGCAGGGAAATGTTAACGGG	240
Db	181	GAAAGCCTGTTGTGTCCACCATCTCCAAGGAGGTTAACTGCAGGGAAATGTTAACGGG	240



```

Query Match      1.9%;      Score 42.8;      DB 4;      Length 72604;
Best Local Similarity 69.4%;      Pred. NO. 0.34;
Matches 75;      Conservative 0;      Mismatches 27;      Indels 6;      Gaps 1;

QY      1      GGAGGTTGAAGTGAGCAGAGATCATGCCA-----GCCTGGGTGACAGTGAGACTCTGTC 54
      |||||
Db      45080      GGAGGTTGCAGTGAGCCGAGATGGTGCCATTGCACTCCAGCCTGACTGACAGAGCAAGAC 45139
      |||||

QY      55      TCAACACAGAATTAAAGGAAAAAAGAAAGAAAAAGAGAGAGAGGAA 102
      |||||
Db      45140      TCCATCTCAAATAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAAA 45187
      |||||

```

Search completed: April 16, 2003, 09:33:58
Job time : 1056 secs

```

FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match          1.9%; Score 43; DB 4; Length 162450;
Best Local Similarity 69.9%; Pred. No. 0.47;
Matches 58; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

y      5 GTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAACAGAA 64
      | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b     54025 GCTGAGATTACACCCTGCACCTCCAGCCTGGGTGACAGTGAGACTCCGAGCTCAAATAA 53966
      | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

y      65 TTAAGGAAAAAGAAAAGAAAGAA 87
      | || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b     53965 TAAATAAATAAATAAAAAATAAAA 53943

RESULT 13
US-09-851-896-3/c
Sequence 3, Application US/09851896
Patent No. 6410325
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Susan M. Freier
APPLICANT: Andrew T. Watt
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)
FILE OF INVENTION: EXPRESSION
FILE REFERENCE: RTS-0220
CURRENT APPLICATION NUMBER: US/09/851,896
CURRENT FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 70000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-851-896-3

Query Match          1.9%; Score 42.8; DB 4; Length 70000;
Best Local Similarity 70.4%; Pred. No. 0.33;
Matches 76; Conservative 0; Mismatches 22; Indels 10; Gaps 1;

QY      1 GGAGGTTGAAGTGAGCAGAGATCATG-----CCAGCCTGGGTGACAGTGAGACTC 50
      | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     6390 GGAGCTTGCAAGTGAGCGAGAGATCATGTCACTGCACCTCCAGCCTGGCGAGAGCGAGACTC 6331
      | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      51 TGTCCTCAAACAGAAATTAAGGAAAAAGAAAGAAAGAAAGAAAGAGAGAGA 98
      | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     6330 TGTCCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 6283

```

LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031

OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57